Identifying Unknown Response Styles: A Latent-Class Bilinear Multinomial Logit Model

Joost van Rosmalen, Hester van Herk and Patrick J.F. Groenen

ERIM REPORT SERIES *RESEARCH IN MANAGEMENT*	
ERIM Report Series reference number	ERS-2007-045-MKT
Publication	July 2007
Number of pages	37
Persistent paper URL	
Email address corresponding author	vanrosmalen@few.eur.nl
Address	Erasmus Research Institute of Management (ERIM) RSM Erasmus University / Erasmus School of Economics Erasmus Universiteit Rotterdam P.O.Box 1738 3000 DR Rotterdam, The Netherlands Phone: + 31 10 408 1182 Fax: + 31 10 408 9640 Email: info@erim.eur.nl Internet: www.erim.eur.nl

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

ABSTRACT AND KEYWORDS	
Abstract	Respondents can vary significantly in the way they use rating scales. Specifically, respondents can exhibit varying degrees of response style, which threatens the validity of the responses. The purpose of this article is to investigate to what extent rating scale responses show response style and substantive content of the item. The authors develop a novel model that accounts for possibly unknown kinds of response styles, content of the items, and background characteristics of respondents. By imposing a bilinear structure on the parameters of a multinomial logit model, the authors can visually distinguish the effects on the response behavior of both the characteristics of a respondent and the content of the item. This approach is combined with finite mixture modeling, so that two separate segmentations of the respondents are obtained: one for response style and one for item content. This latent-class bilinear multinomial logit (LC-BML) model is applied to a cross-national data set. The results show that item content is highly influential in explaining response behavior and reveal the presence of several response styles, including the prominent response styles acquiescence and extreme response style.
Free Keywords	response style, segmentation, visualization, multinomial logit model, cross-cultural research
Availability	The ERIM Report Series is distributed through the following platforms: Academic Repository at Erasmus University (DEAR), DEAR ERIM Series Portal Social Science Research Network (SSRN), SSRN ERIM Series Webpage Research Papers in Economics (REPEC), REPEC ERIM Series Webpage
Classifications	The electronic versions of the papers in the ERIM report Series contain bibliographic metadata by the following classification systems: Library of Congress Classification, (LCC) LCC Webpage Journal of Economic Literature, (JEL), JEL Webpage ACM Computing Classification System CCS Webpage Inspec Classification scheme (ICS), ICS Webpage

Identifying Unknown Response Styles: A Latent-Class Bilinear Multinomial Logit Model

Joost van Rosmalen Hester van Herk Patrick J.F. Groenen*

June 18, 2007

*Joost van Rosmalen is Ph.D. Candidate at Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, Burg. Oudlaan 50, 3000 DR Rotterdam, The Netherlands. Phone: +31 10 4088938. Fax: +31 10 4089162. E-mail: vanrosmalen@few.eur.nl.

Hester van Herk is Associate Professor of Marketing at Department of Marketing, Faculty of Economics and Business Administration, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. Phone: +31 20 5986064. Fax: +31 20 5989870. E-mail: hherk@feweb.vu.nl.

Patrick J. F. Groenen is Professor of Statistics at Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, Burg. Oudlaan 50, 3000 DR Rotterdam, The Netherlands. Phone: +31 10 4081281. Fax: +31 10 4529161. E-mail: groenen@few.eur.nl.

Identifying Unknown Response Styles: A Latent-Class Bilinear Multinomial Logit Model

Abstract

Respondents can vary significantly in the way they use rating scales. Specifically, respondents can exhibit varying degrees of response style, which threatens the validity of the responses. The purpose of this article is to investigate to what extent rating scale responses show response style and substantive content of the item. The authors develop a novel model that accounts for possibly unknown kinds of response styles, content of the items, and background characteristics of respondents. By imposing a bilinear structure on the parameters of a multinomial logit model, the authors can visually distinguish the effects on the response behavior of both the characteristics of a respondent and the content of the item. This approach is combined with finite mixture modeling, so that two separate segmentations of the respondents are obtained: one for response style and one for item content. This latent-class bilinear multinomial logit (LC-BML) model is applied to a cross-national data set. The results show that item content is highly influential in explaining response behavior and reveal the presence of several response styles, including the prominent response styles acquiescence and extreme response style.

Keywords: response style, segmentation, visualization, multinomial logit model, cross-cultural research.

Introduction

Over the last years, the interest in response styles has grown in various disciplines (Baumgartner and Steenkamp 2001; Cheung and Rensvold 2000; De Jong et al. 2007; Van Herk et al. 2004), and response styles are now considered a source of concern in domestic (Greenleaf 1992a; Hui and Triandis 1989) and international research (Johnson et al. 2005). A main reason for this is that response styles threaten the validity of responses given to substantive questions. Furthermore, response styles may be considered real differences between countries by marketing managers and may lead to adverse decisions (Usunier and Lee 2005). Therefore, some researchers have suggested removing response styles from the data (Baumgartner and Steenkamp 2001; Hofstede 2001). However, others state that correcting for response styles might have the effect that content (that is, real differences in opinion among people) is also removed (Fischer 2004); moreover, Smith (2004) even suggested that acquiescence, which is one of the most prominent response styles (Paulhus 1991), is an aspect of the cultural communication style, and that acquiescence as a source of error (bias) should be discounted. Thus, it is important to know the extent to which answers to items reflect a real attitude component or response style.

In addition to separating content from response style, there is the issue of determining the kinds of response styles exhibited by respondents. The majority of articles on response styles focuses on two response styles: acquiescence and/or extreme response style (for example, De Jong et al. 2007; Johnson et al. 2005; Van Herk et al. 2004), but many more response styles have been distinguished (see, for example, Broen and Wirt 1958). Moreover, in the articles on response style, the common procedure for assessing response styles has been the calculation of various response style indices (see Baumgartner and Steenkamp 2001, for an overview). The way in which these response style indices are calculated usually leads to correlated indices; correlations of .50 or higher are often found (see, for example, Baumgartner and Steenkamp 2001; Van Herk et al. 2004).

Furthermore, acquiescence and extreme response style are related to background characteristics of respondents, such as age and education. In his review, Hamilton (1968) already reported accumulated evidence that children and elderly subjects gave more extreme responses than subjects aged 20 to 59. Later studies (for example, Greenleaf 1992a,b) supported this finding. Just as extreme response style, acquiescence seems positively related to age. It was found that older subjects tend to use the positive side of the rating scale more often or display more acquiescence (Greenleaf 1992b); however, others (for example, Johnson et al. 2005) found no relation between age and acquiescence. Education seems negatively associated with acquiescence (Greenleaf 1992a; Krosnick and Alwin 1988; Narayan and Krosnick 1996; Watson 1992) and with extreme response style (Greenleaf 1992a,b). In addition, work in the last decades has shown that there are differences in response styles between cultural populations such as ethnic groups (Bachman and O'Malley 1984; Hui and Triandis 1989; Marin et al. 1992) and countries (Baumgartner and Steenkamp 2001; Chen et al. 1995; De Jong et al. 2007; Van Herk et al. 2004).

The present study contributes to the literature in two ways. First, we provide insight into (a) differences among respondents in reactions to item content (that is, their opinions of the items), (b) differences among respondents in response style (that is, their use of rating scale response categories), and (c) the effects of background characteristics of respondents. Moreover, we visualize how item content affects response behavior, thereby facilitating insight into the extent to which responses to specific items are more likely to get, for example, extreme or acquiescent responses. Second, in contrast to traditional methods using response style indices, we make no a priori assumptions on the kind of the response styles; instead, through segmentation, we can infer what a priori unknown response styles are present in the data. The latter also helps us distinguish between conceptually different response styles (for example, acquiescence and extreme response style) that are often conflated when using response style indices. Although the rating scale categories typically have a natural ordering, we do not impose this ordering on the model, as it could restrict how content and personal characteristics affect response behavior. As Anderson (1984, p. 3) remarks: “There is no merit in fitting an ordered relationship as a routine, simply because the response variable is ordered.” Therefore, we treat the rating categories as unordered in the model.

The outline of this article is as follows. First, we describe the LC-BML model. In an empirical study, we then show how content and response styles can be distinguished by the LC-BML model. Finally, we discuss the implications of our findings, the study’s limitations, and possible topics for further research.

The Latent-Class Bilinear Multinomial Logit Model

In this paper, we seek to model the entire response behavior, that is, the probability that a single respondent ticks certain rating scale categories for the items used in the study, given the content of the items and the respondent’s background characteristics. To do so, we use the responses of all persons to all items in a single statistical model, which we call the Latent-Class Bilinear Multinomial Logit model (LC-BML model). In this model, we distinguish three sets of variables. The first set is a single variable called Rating that indicates the category of the rating scale ticked by the respondent. The second set is the single variable called Item that indicates the item being rated by the respondent. The third set contains variables indicating the background characteristics of the respondents. A nominal measurement level is assumed for all three sets of variables.

The basis of the LC-BML model is an unordered multinomial logit model with Item and the background variables as categorical predictor variables to explain the dependent categorical variable Rating. We adapt this multinomial logit model by incorporating ideas from the models proposed by Anderson (1984), Anderson and Vermunt (2000), and Groenen and Koning (2006). To limit the number of parameters, we apply the bilinear parameter structure used in these articles. In addition, we incorporate latent classes (see, for example, Wedel and Kamakura 2000) to model (a) differences among respondents in reactions to content (that is, their opinions of the items) and (b) differences among respondents in response tendencies (that is, their use of rating scale response categories). Our model extends the stereotype model proposed by Anderson (1984) by using the bilinear parameter structure on categorical predictor variables and by incorporating latent classes. Our approach allows to determine how likely a given pattern of responses is for a single respondent, given the items used in the study and certain characteristics of the respondent.

Before introducing the LC-BML model, we need to define some notation.

i Index to specify the respondent, with $i = 1, \ldots, n$.

t Index to indicate the item being rated, with $t = 1, \ldots, T$.

j Index to indicate the rating, with $j = 1, \ldots, J$.

k Index to indicate the background variable, with $k = 1, \ldots, K$.

Y_{it} The random variable of the rating of person i on item t.

y_{it} The realized rating given by person i on item t.

m_k The number of categories of background variable k.

The multinomial logit model predicts the probability of person i choosing rating j on item t as

$$\Pr(Y_{it} = j) = \frac{\exp(z_{ijt})}{\sum_{j'=1}^{J} \exp(z_{ij't})},$$

with z_{ijt} a linear combination of a constant a_j (the attractiveness of rating category j), the effects of the categories of the background variables, and the effect c_{jt} of rating j on item t, that is,

$$z_{ijt} = a_j + \sum_{k=1}^{K} \mathbf{b}'_{jk}\mathbf{x}_{ik} + c_{jt}. \tag{1}$$

Each background variable k can be represented by a matrix $\mathbf{X}_k$ of dummy variables for each category so that $\mathbf{x}'_{ik}$ is row i of $\mathbf{X}_k$, and $\mathbf{b}'_{jk}\mathbf{x}_{ik}$ selects the element from vector $\mathbf{b}_{jk}$ belonging to the category of respondent i on background variable k. As a result of Equation 1, separate model parameters need to be estimated for each rating scale category. To see this, consider a hypothetical example with $J = 3$ response categories of Rating, $T = 4$ items, and $K = 1$ background variable (that has $m_1 = 2$ categories). For this case, Table 1 contains the z_{ijt}s as the linear combination of the specific parameters. There are $J = 3$ attractiveness parameters a_j, $J \times m_1 = 6$ parameters for the effects of the two categories of the background variable, and $J \times T = 12$ parameters for the effect c_{jt} of item t on rating j.

To fully model the effects of the explanatory variables on the dependent variable, a parameter vector is required for each category of each explanatory variable. The number of parameters $J + J\sum_{k=1}^{K} m_k + JT$ gets large if the total number of categories is large. The resulting parameter vectors are generally not easy to represent or to interpret.

The problem of the large number of parameters occurs in a wide variety of models for analyzing categorical data sets. For the case of log-linear analysis, Anderson and Vermunt (2000) proposed several models that construct a parsimonious representation of the relationships between categorical variables, based on a bilinear decomposition. Groenen and Koning

Table 1: The linear combination of the z_{ijt}s for a hypothetical example with $J = 3$ response categories of Rating, $T = 4$ items, and $K = 1$ background variable.

Rating j	Item 1	Item 2	Item 3	Item 4
1	$a_1 + \mathbf{b}'_1\mathbf{x}_{i1} + c_{11}$	$a_1 + \mathbf{b}'_1\mathbf{x}_{i1} + c_{12}$	$a_1 + \mathbf{b}'_1\mathbf{x}_{i1} + c_{13}$	$a_1 + \mathbf{b}'_1\mathbf{x}_{i1} + c_{14}$
2	$a_2 + \mathbf{b}'_2\mathbf{x}_{i1} + c_{21}$	$a_2 + \mathbf{b}'_2\mathbf{x}_{i1} + c_{22}$	$a_2 + \mathbf{b}'_2\mathbf{x}_{i1} + c_{23}$	$a_2 + \mathbf{b}'_2\mathbf{x}_{i1} + c_{24}$
3	$a_3 + \mathbf{b}'_3\mathbf{x}_{i1} + c_{31}$	$a_3 + \mathbf{b}'_3\mathbf{x}_{i1} + c_{32}$	$a_3 + \mathbf{b}'_3\mathbf{x}_{i1} + c_{33}$	$a_3 + \mathbf{b}'_3\mathbf{x}_{i1} + c_{34}$

(2006) proposed a similar model in the context of analysis of variance, with the aim of obtaining a simple low-dimensional visual display of the estimated interaction effects. Based on their interaction decomposition model, we adapt the multinomial logit model described above to parsimoniously represent the effects of the explanatory variables. An important objective of our approach is the graphical representation of the parameter estimates, to facilitate a better understanding of the effects.

To illustrate this approach, consider the hypothetical example given in Figure 1a. In this figure, there are again three response categories of Rating, four items, and one background variable with two categories. Each rating category is represented by a vector reaching from the origin outwards. Each category of each explanatory variable (Item or the background variable) is represented by a dot. The effects of the explanatory variables are related to the projections of the dots onto the vectors of Rating. For example, if one is interested in the estimated effect of Item 4 on Rating 1, one may proceed as follows.

1. Project the point 'Item 4' onto the vector 'Rating 1'.

2. Multiply the length of the projection with the length of the vector 'Rating 1'.

3. The resulting length is the estimated effect, which also equals the inner product of the point and the vector.

In Figure 1a, all points of Item and the background variable have been projected onto Rating 1. In the model, the projected effects are multiplied by the length of the vector Rating 1, and the result is shown in Figure 1b. Here, one can see that Item 4 has a relatively high estimated effect on Rating 1, whereas the effects of Items 2 and 3 are clearly negative. The estimated effects of the background variable are relatively small, as all its projections are close to zero.

This representation has the following properties. Long vectors and points far away from the origin tend to correspond to large effects. In addition, if the angle between the vector

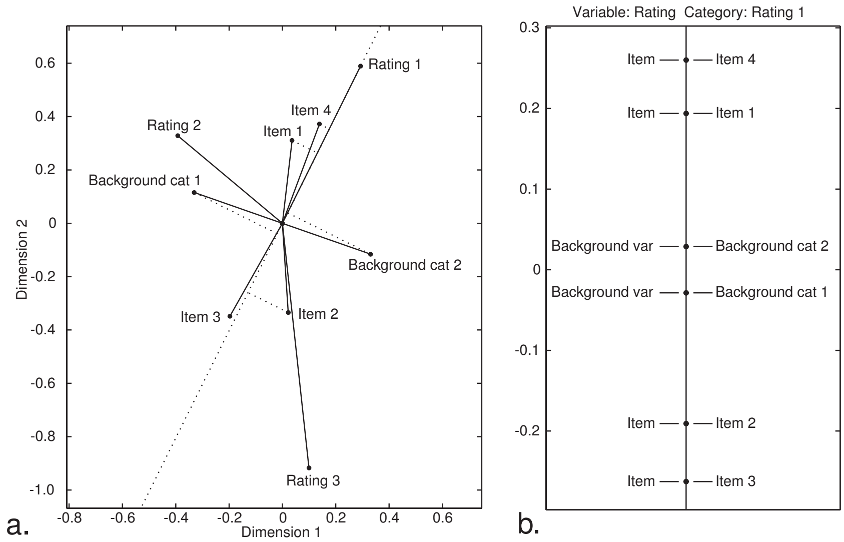

Figure 1: Hypothetical example of a graphical representation in the bilinear multinomial logit model. Panel a shows the graphical representation with the ratings as vectors, the categories of the predictor variables as points, and the projections of the points of the categories of the predictor variables on Rating 1. Panel b shows the projections multiplied by the length of Rating 1.

and the line segment between the point and the origin is smaller than 90 degrees, the corresponding estimated effect is positive. If this angle is greater than 90 degrees, the estimated effect is negative. This effect becomes stronger as the angle gets closer to either 0 or 180 degrees.

We now proceed with a more formal description of the bilinear decomposition part of the model. Remember that $\mathbf{b}_{jk}$ is the effect of background variable k on rating j. Let us gather the $\mathbf{b}'_{jk}$s for all ratings j underneath each other, that is, by $\mathbf{B}_k = [\mathbf{b}_{1k}|\mathbf{b}_{2k}|\cdots|\mathbf{b}_{Jk}]'$. To ensure that the rating scale categories are represented by vectors and the categories of the background variables and the items by points, we restrict each $\mathbf{B}_k$ and $\mathbf{C}$ (with elements c_{jt}) as

$$\mathbf{B}_k = \mathbf{F}\mathbf{G}'_k \text{ and } \mathbf{C} = \mathbf{F}\mathbf{H}'. \tag{2}$$

Here, $\mathbf{F}$ is the $J \times P$ matrix of coordinates for the vectors of ratings in the visual representation, $\mathbf{G}_k$ the $m_k \times P$ matrix of coordinates for background variable k, and the $T \times P$ matrix $\mathbf{H}$ denotes the coordinates of the T items, where P is the dimensionality of the representation. The model under the rank restrictions in Equation 2 can now be described as

$$z_{ijt} = a_j + \sum_{k=1}^{K} \sum_{p=1}^{P} f_{jp} \mathbf{g}'_{kp} \mathbf{x}_{ik} + \sum_{p=1}^{P} f_{jp} h_{tp},$$

where f_{jp} and h_{tp} are elements of $\mathbf{F}$ and $\mathbf{H}$ respectively, and $\mathbf{g}_{kp}$ is the p-th column of $\mathbf{G}_k$.

Segmentation of Respondents

In marketing, response behavior cannot be considered homogeneous across respondents. People differ with respect to their opinions and preferences, and heterogeneity in response behavior is expected. So far, the model has no possibility to take heterogeneity in the response behavior of respondents into account, and assumes that the T responses of every respondent are independent. Depending on what method is used, modeling these effects may require a large number of additional parameters. To account for dependent observations and unobserved heterogeneity in the respondents, we use finite mixture modeling (see, for example, Kamakura and Russell 1989). We extend our model to take two types of heterogeneity into account. First, respondents may vary with respect to their response styles, so that they differ in the probabilities with which they generally tick different response categories of Rating. We account for these differences in response styles by allowing the constants a_j to vary between latent classes of respondents. Second, respondents can differ on their opinions of the relative importance of the items. These differences are incorporated by allowing the parameters h_{tp} to also vary between latent classes of respondents. We thus construct a joint segmentation of the respondents in the model, in which response style and content function as bases for the two types of segmentations. We assume that the prior probabilities of membership in the two types of segments are independent of each other.

Let there be R segments for the response tendencies with mixing proportions u_r, so that $\sum_{r=1}^{R} u_r = 1$. Also, let S be the number of segments for representing the opinions of the relative importance of the items with mixing proportions v_s and restriction $\sum_{s=1}^{S} v_s = 1$. We refer to these types of segments as response style segments and item segments, respectively. The probability that respondent i evaluates item t with rating j can be calculated as

$$\Pr(Y_{it} = j) = \sum_{r=1}^{R} \sum_{s=1}^{S} u_r v_s \Pr(Y_{it} = j|r, s), \tag{3}$$

with

$$\Pr(Y_{it} = j|r, s) = \frac{\exp(z_{ijt|r,s})}{\sum_{j'=1}^{J} \exp(z_{ij't|r,s})}, \tag{4}$$

$$z_{ijt|r,s} = a_{j|r} + \sum_{k=1}^{K} \sum_{p=1}^{P} f_{jp} \mathbf{g}'_{kp} \mathbf{x}_{ik} + \sum_{p=1}^{P} f_{jp} h_{tp|s}.$$

The likelihood of observing responses y_{it}, $t = 1, \ldots, T$, conditional on respondent i belonging to response style segment r and item segment s is

$$L_i(\mathbf{y}_i|r, s) = \prod_{t=1}^{T} \prod_{j=1}^{J} \Pr(Y_{it} = j|r, s)^{I(y_{it}=j)},$$

where $I()$ denotes the indicator function, and $\mathbf{y}_i$ is the vector of length T with the observed responses of respondent i. As the unconditional likelihood for respondent i is

$$L(\mathbf{y}_i) = \sum_{r=1}^{R} \sum_{s=1}^{S} u_r v_s L_i(\mathbf{y}_i|r, s),$$

the likelihood of the entire model, as a function of the model parameters, is given by

$$\begin{aligned} & L(a_{.|1}, \ldots, a_{.|R}, \mathbf{F}, \mathbf{G}_1, \ldots, \mathbf{G}_K, \mathbf{H}_1, \ldots, \mathbf{H}_S, \mathbf{u}, \mathbf{v}) \\ & = \prod_{i=1}^{n} \sum_{r=1}^{R} \sum_{s=1}^{S} u_r s_v \prod_{t=1}^{T} \prod_{j=1}^{J} \Pr(Y_{it} = j | r, s)^{I(y_{it}=j)}. \end{aligned} \tag{5}$$

A number of parameter constraints are required to ensure parameter identification of the LC-BML model; these constraints are described in Appendix A. For given values of the numbers of segments R and S and the dimensionality P, parameter estimates are obtained by maximizing the likelihood function in Equation 5. A description of the optimization algorithm can be found in Appendix B. Given the optimal parameter estimates, the posterior probability that respondent i belongs to response style segment r and item segment s can be calculated in a Bayesian way as

$$\Pr(i \in \Xi_{r,s}) = \frac{u_r v_s \prod_{t=1}^{T} \prod_{j=1}^{J} \Pr(Y_{it} = j | r, s)^{I(y_{it}=j)}}{\sum_{r'=1}^{R} \sum_{s'=1}^{S} u_{r'} v_{s'} \prod_{t=1}^{T} \prod_{j=1}^{J} \Pr(Y_{it} = j | r', s')^{I(y_{it}=j)}}, \tag{6}$$

where $\Xi_{r,s}$ denotes the set of respondents belonging to response style segment r and item segment s.

The numbers of segments are typically determined using an information criterion, such as AIC or BIC (see Andrews and Currim 2003, for an overview of the performance of these criteria). In this paper, we will use the consistent Akaike information criterion (CAIC) to help determine the numbers of segments R and S. This criterion performs well with very large sample sizes (see Andrews and Currim 2003; Bozdogan 1987) and has a low risk of overfitting.

Data

The data used in this study are based on a commercial survey performed in 1996 (see also Van Herk 2000). The sample used consisted of 3840 male respondents from five European countries (France, Italy, Germany, the United Kingdom, and Spain). For each respondent, the background variables Country, Age, and Education were measured. The numbers of respondents were 823, 994, 764, 698, and 561 in France, Italy, Germany, the United Kingdom, and Spain, respectively. Age was coded as a categorical variable with response categories $15-24$ (13%), $25-34$ (15%), $35-44$ (16%) , $45-54$ (21%), $55-64$ (22%), and 65+ (13%). Education had two levels: higher level of education (42%) and lower level of education (58%). The items that were used in this study are on the List of Values (Kahle 1983), which has been widely used in marketing (see, for example, Brunso et al. 2004; Kamakura and Novak 1992; Madrigal and Kahle 1994). It comprises nine items, of which respondents indicate the importance. The nine items are (a) sense of belonging, (b) excitement, (c) warm relationships with others, (d) self-fulfillment, (e) being well-respected, (f) fun & enjoyment in life, (g) security, (h) self-respect, and (i) a sense of accomplishment. Each item was measured on a 9-point rating scale ranging from '1' ('very important') to '9' ('very unimportant').

The procedure used for assessing translation equivalence was parallel translation (Craig and Douglas 2000). Bilinguals translated all questionnaire items from English into Italian, French, Spanish, and German. Next, an independent group of researchers criticized the translations. As a last step, a discussion between the project coordinators in the participating countries was held to choose the version that most closely resembled the original English version. The List of Values has been employed in single country studies (see, for example, Brunso et al. 2004; Kamakura and Novak 1992) as well as in studies in multiple countries (Grunert et al. 1989; Soutar et al. 1999; Wedel et al. 1998).

Results

In this section, we apply the LC-BML model to the List of Values data set and interpret the results. We seek to explain the observed rating values using the LOV item and the background characteristics Country, Age, and Education as predictor variables. In our analyses, we use each respondent's answers on how important each LOV item is in his daily life, so that we have $3840 \times 9 = 34560$ responses in total. Before we can interpret the results, the dimensionality P, the number of response style segments R, and the number of item segments S need to be chosen. We choose both R and S using the consistent Akaike information criterion (CAIC). We set $P = 2$, as the results show that the first two dimensions of the graphical representations can be interpreted in terms of known response styles. Solutions with $P = 3$ show that a third dimension would not allow for such an interpretation. We estimated the model parameters for every combination of $R = 1, \ldots, 13$ response style segments and $S = 1, \ldots, 7$ item segments. The resulting CAIC values are shown in Table 2. The lowest CAIC value $(110, 156)$ is attained with 11 response styles segments and 4 item segments; the corresponding value of the log-likelihood is $-53, 950$. We use these numbers of segments to interpret the model results for the LOV data set.

In the first part of our analysis, we show the values of the parameters that are used for the bilinear decomposition part of the model ($\mathbf{F}$, $\mathbf{G}_1, \ldots, \mathbf{G}_K$, and $\mathbf{H}_1, \ldots, \mathbf{H}_S$), using a number of graphical representations similar to Figure 1a. Figure 2 shows the effects of the background variables on Rating, which are the same for each item segment. In this figure, the points of the background variables (which are contained in $\mathbf{G}_1, \ldots, \mathbf{G}_K$) are shown as dots and the points belonging to Rating (which are contained in $\mathbf{F}$) are shown as vectors. From this figure, we can see that the structure of the categories of Rating resembles a U-shape. The categories of Rating have retained their natural ordering of 1 through 9 with respect to the first dimension.[1] Therefore, respondents with background characteristics with a high score

[1]However, one should be careful in assigning interpretations to the dimensions, as the graphical represen-

Table 2: CAIC values with $P = 2$ for $R = 1, \ldots, 13$ response style segments and $S = 1, \ldots, 7$ item segments in the LOV data set

Response style	Item segments						
segments	1	2	3	4	5	6	7
1	120,273	118,467	117,137	116,753	116,375	116,077	116,037
2	116,201	115,444	114,383	113,368	113,441	113,128	112,949
3	113,489	112,779	112,557	112,403	112,164	112,022	112,090
4	112,555	111,779	111,433	111,426	111,246	111,217	111,164
5	112,209	111,339	111,097	110,981	110,876	110,866	110,864
6	111,944	111,062	110,809	110,659	110,599	110,626	110,635
7	111,729	110,850	110,593	110,456	110,412	110,421	110,364
8	111,644	110,738	110,459	110,286	110,302	110,307	110,327
9	111,614	110,688	110,383	110,197	110,220	110,253	110,288
10	111,582	110,647	110,350	110,161	110,190	110,241	110,278
11	111,593	110,652	110,343	110,156	110,184	110,243	110,267
12	111,600	110,669	110,352	110,159	110,181	110,241	110,270
13	111,617	110,670	110,369	110,174	110,191	110,256	110,282

on the first dimension tend to tick the rating category ‘1’ (that is, ‘very important’) relatively frequently. Moreover, the scores of the background characteristics on the first dimension are positively related to the probability of ticking positive rating categories such as ‘1’ or ‘2’. For example, respondents between 15 and 24 years of age tend to tick the ‘1’ and ‘2’ more often than people in the other age categories. On the second dimension, the most extreme rating categories ‘1’ and ‘9’ have the highest scores, whereas moderate responses such as ‘3’, ‘4’, and ‘5’ have low scores. Therefore, we can infer that the values of the background characteristics on the second dimension are positively related to the probability of choosing extreme response categories. It can be seen, for example, that elderly people and people from Italy tend to use the extreme response categories more often than other people.

In short, Figure 2 confirms the expectation that respondents with a low education level are more prone to ticking extreme response categories than respondents with a high level of education. Further, age appears negatively related to the probability of ticking the positive side of the rating scale and seems to have a curvilinear relationship with the probability

tations can, in principle, be rotated arbitrarily; here, they are rotated so that the greatest variation of the coordinates occurs along the first dimension, see Appendix A.

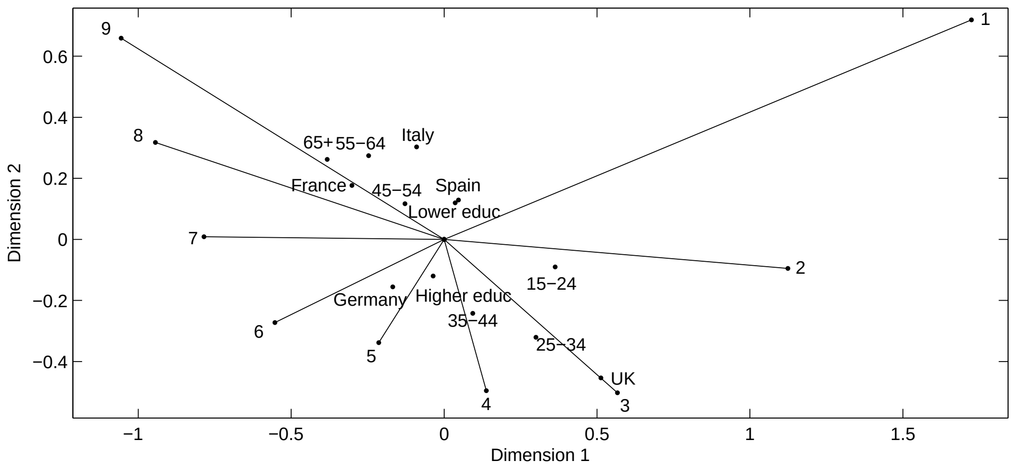

Figure 2: Graphical representation of background characteristics in LOV data set

of ticking extreme response categories. Moreover, the respondents in Southern European countries show a higher tendency to tick extreme response categories than respondents in the U.K. and Germany.

Item Segments

To show how content differs across the item segments, we provide a separate plot of the effects of Item on Rating (using the parameters $\mathbf{H}_s$), for each item segment s. In Figure 3, the points of the items are shown as dots with labels, the background variables as dots without labels, and the rating categories are again shown using vectors. In every item segment, the effects of Item on the response behavior tend to be larger than the effects of the background variables, as the points belonging to Item are generally farther away from the origin. In addition, the differences in content between the item segments tend to be quite large. For example, 72.8% of the respondents in item segment 1 evaluated 'belonging' with a rating of '1', whereas this percentage was only 4.8 for the respondents in item segment 4. In contrast, for the item 'respected', these percentages were 9.3 and 48.8, respectively.

To further clarify the magnitude of the differences between the item segments, Table

3 shows the average rating value per item, for both the item segments and the response style segments. These average rating values are calculated using the posterior segment membership probabilities in Equation 6. In the item segments part of Table 2, it can be seen that respondents in segments 1, 2, and 3 tend to distinguish more between the different items than respondents in segment 4. For example, in segment 3, the average ratings range from 1.83 ('security') to 5.78 ('excitement'), whereas the range in segment 4 is from 2.74 ('belonging') to 3.54 ('respected'). These results show that respondents in the different segments differ with respect to the importance they attach to specific items (for example, is 'respected' more or less important than 'belonging') as well as in the rating score they attach to the item (that is, do they tick a rating score of about '3' or use other rating scores as well).

In the response style segments part of Table 2, it can be seen that respondents in response style segment 1 have an average score of 4.01 for belonging. Response style segment 2 shows averages of less that 2.00 for all items except 'belonging' and 'excitement', implying that those people consider 7 out of 9 LOV items very important. In contrast, all averages in response style segment 11 are above 5.70; this might imply that these respondents consider all 9 LOV items quite unimportant.

Item Segments and Respondent Characteristics

Next, we show what types of respondents are present in each segment. Table 4 shows what percentage of the respondents from each category of each background variable belongs to a certain segment, for both the item segments and the response style segments. These percentages are also calculated using the posterior probabilities that a respondent belongs to a certain segment in Equation 6. The final row shows the segment sizes, which are equal to the values of v_s for the item segments and the values of u_r for the response style segments. As the parameters $\mathbf{G}_1, \ldots, \mathbf{G}_K$ model the effects of the background variables on the response styles, the distribution of the respondents over the response style segments should be approximately equal for all categories of the background variables. The right-hand side of Table 4 shows that this requirement is met to a reasonable degree, as the differences between the rows are fairly small. For the item segments, these differences are certainly not small, nor do they need to be. The differences in content between countries appear to be particularly large.

Using the information in both Figure 3 and Table 4, we can further interpret the item segments. In the largest item segment 1 (39%), respondents attach high importance to the LOV-items 'fun & enjoyment', 'warm relationships with others', and 'being well respected'. The values 'sense of belonging' and 'excitement' both are considered far less important. In item segment 2 (26%), 'self-respect' and 'being well respected' are considered the most important. The values 'fun & enjoyment', 'warm relationships', and especially 'belonging' are considered less important. This segment is mainly found in Italy, followed by the UK and Spain. Respondents in item segment 3 (19%) consider 'belonging', 'self-respect', and 'security' the most important values; 'being well respected', 'self-fulfillment', and 'excitement' are considered less important. In the smallest item segment 4 (16%), no LOV item received low scores. The most important values for this segment are 'belonging' and 'fun & enjoyment'. For both 'self-respect' and 'fun & enjoyment', relatively more extreme response categories are ticked; 'excitement' is often considered of average importance.

The above description of the four item segments indeed shows that the segments include respondents who have different opinions on the importance of the LOV items. Moreover, the segments are quite similar to segments based on the LOV that are found in the literature. For example, in item segment 1 there is a relatively high percentage of young (that is, younger than 25 years of age) respondents. In a study by Kamakura and Novak (1992), it was also found that value priorities as in item segment 1 (for example, considering 'fun & enjoyment' and 'warm relationships' important) apply to younger respondents. In addition, the ordering

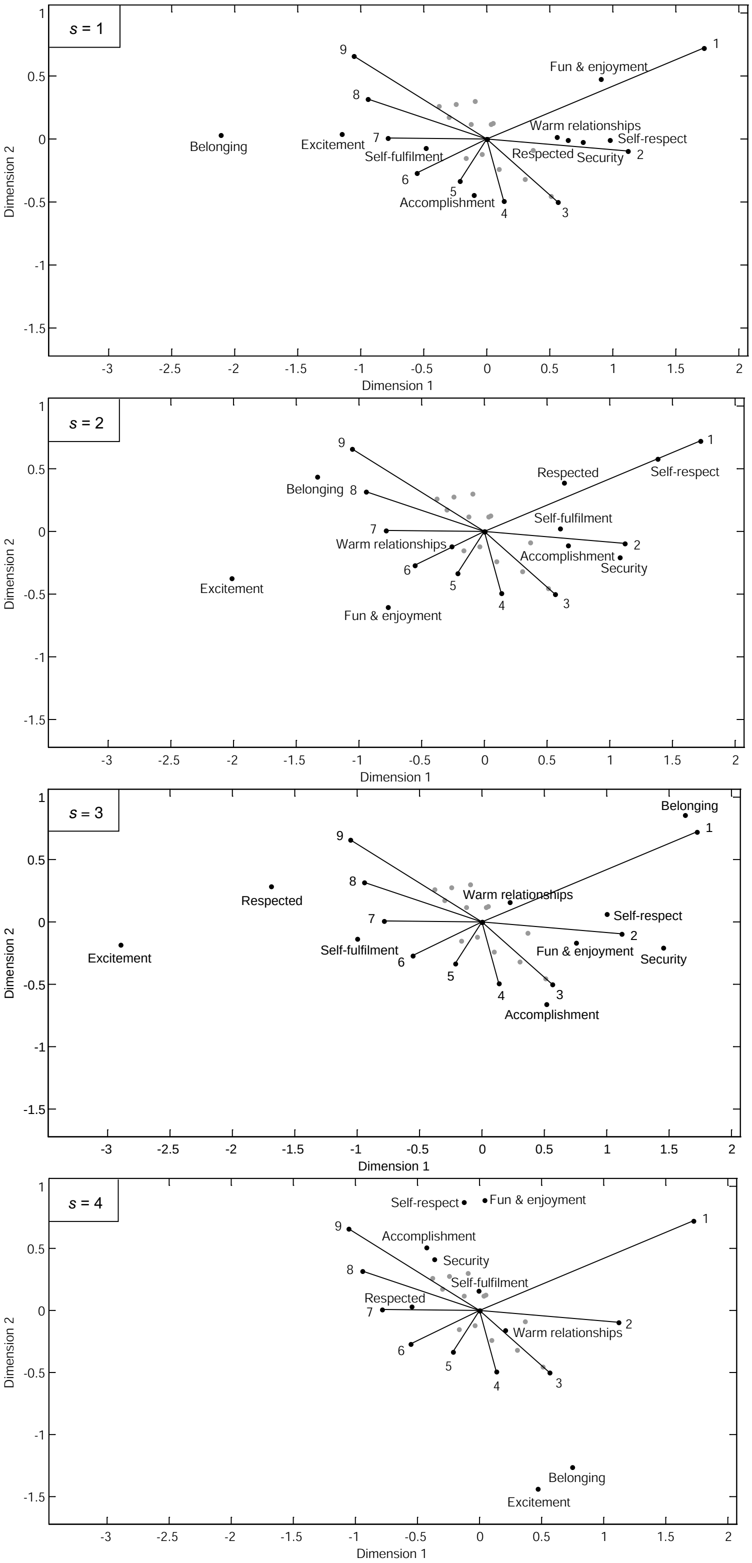

Figure 3: Graphical representation of item segments $s = 1$ through 4 in LOV data set. The categories of the background variables are not labeled in these plots.

Table 3: Average Rating Score for Every Item, in Every Segment of the LOV Data Set

	Item segment s				Response style segment r											
Item	1	2	3	4	1	2	3	4	5	6	7	8	9	10	11	Total
Belonging	5.18	4.09	1.49	2.74	4.01	2.46	3.47	3.32	5.28	4.39	4.49	5.12	6.41	6.59	5.89	3.81
Excitement	3.99	4.90	5.78	2.99	4.69	3.03	4.02	3.86	6.28	4.76	5.24	5.52	6.81	7.04	6.15	4.41
Warm relationships	2.20	2.85	2.50	2.86	2.43	1.62	1.70	2.46	4.05	3.49	2.65	3.03	3.98	7.82	5.73	2.53
Self-fulfilment	3.21	2.07	3.82	2.96	3.08	1.84	2.01	2.77	4.92	3.96	3.08	3.87	5.15	7.90	5.83	2.99
Respected	2.14	1.94	4.67	3.54	2.80	1.78	1.90	2.67	4.53	3.78	2.81	3.06	4.04	8.02	5.95	2.79
Fun & enjoyment	1.85	3.52	2.21	2.75	2.39	1.61	1.72	2.42	3.95	3.39	2.63	2.65	3.38	8.22	6.04	2.49
Security	2.07	1.86	1.83	3.26	1.98	1.39	1.22	2.26	3.42	3.29	2.15	2.19	2.86	8.09	6.15	2.16
Self-respect	1.94	1.55	1.98	2.92	1.75	1.27	1.11	2.03	3.28	3.21	1.91	2.12	2.34	8.29	6.36	2.00
Accomplishment	2.91	2.07	2.54	3.33	2.68	1.69	1.53	2.54	4.71	3.69	2.62	3.81	4.25	8.19	6.17	2.69

of LOV-items in segment 3 is typical for German respondents (see, for example, Grunert and Scherhorn 1990; Wedel et al. 1998), and our data also show an overrepresentation of German respondents (64%) in this segment. Moreover, as found by Kamakura and Novak (1992), the respondents who consider 'belonging' and 'security' relatively important (item segment 3) often are older.

Response Style Segments

Furthermore, we show what response styles are used by the respondents in each response style segment. Table 5 shows the relative frequencies of the rating categories in percentages for the response style segments. This table contains the response profile of each segment. Instead of interpreting the parameters $a_{j|r}$, we show the relative frequencies of ratings per segment in Table 5, as these frequencies more directly show what rating categories have been used by the respondents in every segment. The 11 response style segments identified by our model fit into the possible response styles distinguished by Broen and Wirt (1958).

For instance, there is a segment in which respondents tend to spread ratings (segment 5); there are segments in which respondents mostly tick positive rating categories (segments 2 and 3) and segments in which respondents tick negative rating categories (segments 10 and 11). Further, a segment is found in which respondents tend to go to extremes (segment 9), and two segments are found in which there is a tendency to avoid extreme ratings (segments 6 and 11). Interestingly, respondents in response style segment 8, in addition to extreme responding, also show a high degree of midpoint scoring. The respondents in response style segment 10 seem to have an extremely high level of disacquiescence, as Table 5 shows that over 80% of their responses are either '7', '8', or '9'. However, a closer inspection of Table 3 shows that values that are rated as relatively important by most respondents are considered relatively unimportant by the respondents in segment 10, and vice versa. Therefore, we conclude that the respondents in segment 10 typically must have misread the rating scale

Table 4: Segment Membership Percentages for Each Category of Each Background Variable in LOV Data Set

	Item Segment s				Response Style Segment r										
	1	2	3	4	1	2	3	4	5	6	7	8	9	10	11
Country															
France	69.9	12.5	6.1	11.4	24.2	19.9	16.9	12.1	7.4	6.2	4.9	4.0	2.7	.6	1.0
Italy	31.6	49.0	3.7	15.7	19.7	22.0	18.8	11.1	7.8	6.9	4.4	2.7	3.1	1.3	2.2
Germany	10.8	8.2	64.0	17.0	25.0	18.7	16.5	15.8	7.6	6.9	4.7	2.0	2.0	.3	.6
UK	29.7	29.3	19.5	21.4	20.3	20.8	16.7	11.8	7.4	8.8	3.1	2.6	.9	5.5	2.1
Spain	57.5	26.7	3.1	12.7	25.3	17.9	21.2	11.5	6.0	6.4	4.0	3.0	1.7	.9	2.0
Education															
Lower educ	36.7	26.5	21.9	15.0	22.4	19.9	18.5	12.2	7.3	7.2	4.2	3.0	2.3	1.5	1.6
High Educ	42.5	25.9	15.0	16.6	23.0	20.4	17.1	12.8	7.4	6.8	4.3	2.6	2.0	1.8	1.6
Age															
25-	47.8	19.7	10.4	22.1	22.3	22.7	18.2	11.6	7.6	6.3	2.8	2.6	2.3	1.7	1.9
25-34	44.5	20.1	15.2	20.3	21.8	20.6	17.4	12.1	7.7	7.3	3.9	3.1	2.5	2.0	1.6
35-44	40.2	25.8	19.5	14.4	24.5	17.8	18.2	12.9	6.6	7.1	4.5	2.8	1.7	1.8	2.0
45-54	31.2	29.7	25.2	14.0	23.2	19.7	17.2	11.4	7.6	7.6	4.8	3.3	1.8	1.9	1.5
55-64	34.3	32.3	22.2	11.2	21.8	20.9	18.1	13.2	7.8	6.4	4.7	2.2	3.1	.8	1.0
65+	34.7	33.1	21.8	10.4	21.6	20.0	18.8	13.4	6.9	7.3	4.7	2.8	1.9	1.2	1.4
Segment size (v_s, u_r)	39.1	26.2	19.0	15.7	22.6	20.1	17.9	12.4	7.3	7.0	4.3	2.8	2.2	1.6	1.6

Table 5: Relative Frequencies of Rating Values in Percentages and Response Style Indices for Each Response Style Segment in LOV Data Set

	Response Style Segment r											
Rating	1	2	3	4	5	6	7	8	9	10	11	Overall
1	34.7	51.3	70.2	13.2	16.2	4.2	13.0	45.7	48.1	.6	.6	36.8
2	17.8	28.8	4.9	39.8	13.8	12.3	51.3	1.8	4.1	.0	1.0	19.9
3	17.3	11.6	6.3	25.8	13.0	32.2	8.9	4.3	3.0	.9	2.9	14.4
4	9.8	4.2	4.8	12.1	9.4	23.7	3.7	3.1	2.8	1.3	7.3	8.2
5	9.7	2.1	6.3	4.9	11.8	16.4	11.6	31.0	2.1	6.0	24.8	8.3
6	3.2	.5	1.8	2.4	7.9	6.1	2.1	2.2	1.2	6.5	25.8	3.2
7	3.6	.5	1.3	1.2	9.8	3.3	1.9	.9	2.4	19.4	22.1	3.1
8	2.0	.3	1.1	.5	9.2	1.5	2.4	.0	1.9	21.8	10.6	2.2
9	1.8	.6	3.3	.0	8.9	.4	5.1	11.0	34.4	43.5	4.8	3.9
Acquiescence index	52.5	80.1	75.2	53.0	30.0	16.4	64.3	47.5	52.2	.6	1.6	56.7
ERS index	36.5	51.9	73.5	13.3	25.1	4.6	18.0	56.7	82.4	44.1	5.4	40.7

and incorrectly thought that ‘9’ was the most positive rating value. The percentage of respondents who have done this ‘scale inversion’ is relatively small (the size of segment 10 is 1.6%), but due to the severity of the mistake, it can have major consequences for the results of a study; this mistake could also have been made by some respondents in segment 11.

To gain insight into the validity of our model, the acquiescence index and the ERS index, calculated as in Bachman and O’Malley (1984), are given in the last two rows of Table 5. The differences among the segments in these response style indices show that there are large variations in both the level of acquiescence and the level of extreme response style. In addition, a high level of acquiescence does not necessarily imply a high level of extreme response style or vice versa. For example, the acquiescence index is approximately equal for segments 1, 4, and 9, but the ERS index takes values of 36.5%, 13.3%, and 82.4% for these segments, respectively. Although segments 1 and 9 have almost identical acquiescence indices, Table 5 shows that the response style of segment 1 is clearly more acquiescent than the response style of segment 9; therefore, the response style indices may not always adequately capture the response style they are supposed to measure. These results show that the LC-BML model can distinguish between response styles such as acquiescence and

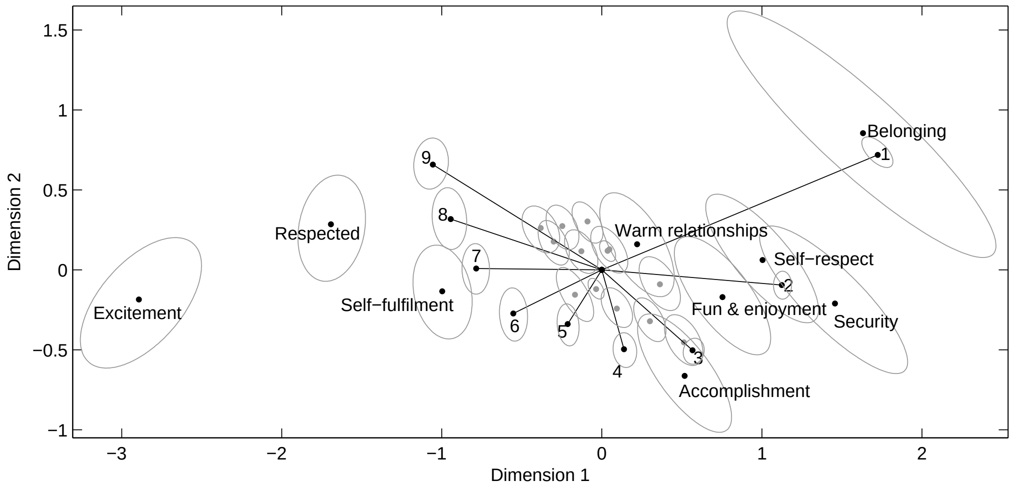

Figure 4: Graphical representation of item segment 3 with 95% confidence ellipses for all points.

ERS, which are often conflated when using response style indices.

The response style segments reveal several response styles mentioned in the literature such as acquiescence (most prominently in response style segments 2 and 3), extreme response style (response style segment 9), and midpoint scoring (response style segment 8). However, a tendency to tick rating categories that are moderately positive, but not extremely positive, is found for most respondents in our data set (segments 1 and 4 through 7). Such a moderately positive answering style is not considered a response style or response bias in the literature; however, it seems an important way of ticking rating categories.

Reliability of Results

To show the statistical significance of our results, we have constructed confidence ellipses for the points in our graphical representations. Figure 4 shows 95%-confidence ellipses for all points shown in the graphical representation of item segment 3. An explanation of how the confidence ellipses are constructed is given in Appendix C. The amount of uncertainty in the graphical representation appears to be low in general. The confidence ellipses are larger for the items than they are for the ratings and the background variables, as only the respondents in item segment 3 determine the locations of the points of the items in this plot. The locations of all points are significantly different from the origin, so that the effects of all predictor variables are statistically significant. The confidence ellipses of Rating do not overlap with each other, which means that all rating categories are distinguishable with respect to the predictor variables (see Anderson 1984, for more information on the distinguishability of rating categories).

Conclusions and Discussion

In this article, we proposed a new model, called the LC-BML model, for explaining response behavior and distinguishing between content and response styles in data sets with rating scale responses. Unlike other methods to assess response style, that either need a scale especially developed to measure a specific response style (Greenleaf 1992a) or a large set of heterogeneous items (see, for example, Baumgartner and Steenkamp 2001), the LC-BML model can assess response style in a data set with a limited number of variables. Specifically, we found that there are several different response styles as suggested earlier by Broen and Wirt (1958). The most prominent response style is acquiescence (38.0% of respondents), followed by spreading ratings evenly across all rating scale categories (7.3%), avoiding extremes (7.0%), using extremes (5.0%), and mainly ticking negative categories (3.2%). The remaining respondents (39.3%) do show a response style in which they use several response categories, but do not tick the extremely high or low values more often. These respondents mainly tick moderately positive categories ('2' and '3'). These results suggest that, in studies on values, acquiescence is the main threat to the validity of the findings.

The LC-BML model distinguishes between the effects of the content of an item, background characteristics, and the response style of a respondent, which contributes to our understanding of response behavior in the following ways. First, the model shows that item content is highly influential in explaining differences in response behavior; there are large differences among the effects of the items on response behavior within each item segment, and the relative importance of the items also varies greatly between the item segments. Second, the effects of the background characteristics (that is, country, age, and education) on response behavior are much smaller than the effects of the content of the items. Third, the LC-BML model reveals what response styles are used by the respondents. Unlike studies using response style indices, which predefine specific response styles, the LC-BML model does not impose specific response styles. This allowed us, for example, to infer that some respondents reversed the scale; they considered '9' instead of '1' very important, so that their answers were mostly negative. In a study in which response style indices are used, these respondents would be classified as people showing disacquiescence. Our study shows that this inference would have been incorrect and helps researchers decide on removal of these inappropriate scores from the data. Fourth, the LC-BML model, unlike other models for response style, in which response styles are often conflated, enables us to distinguish between acquiescence and extreme response style. Our results show, for example, that respondents in Southern European countries tend to give more extreme responses, but they do not appear to acquiesce more often.

Because the LC-BML model accounts for differences in both response styles and opinions, it can also be used to correct for the presence of response styles (that is, remove the response styles from the data). For each observation, we can determine estimated posterior response probabilities as

$$\Pr(Y_{it} = j|\mathbf{y}_i) = \sum_{r=1}^{R} \sum_{s=1}^{S} \Pr(i \in \Xi_{r,s}) \Pr(Y_{it} = j|r, s), \tag{7}$$

where $\Pr(i \in \Xi_{r,s})$ and $\Pr(Y_{it} = j|r, s)$ are calculated using Equations 4 and 6, respectively. These estimated posterior response probabilities can be corrected for the presence of response styles as follows. First, correct the conditional response probabilities $\Pr(Y_{it} = j|r, s)$ by setting $a_{j|r} = 0$ for all j and r and by setting all elements of each estimated $\mathbf{g}_{kp}$ equal to 0 in Equation 4. Then, recalculate the posterior response probabilities using Equation 7. Note that the original (uncorrected) parameter estimates should be used to compute the posterior segment membership probabilities $\Pr(i \in \Xi_{r,s})$. The corrected response probabilities may be used to simulate a data set from which differences in response styles have been removed as much as possible. Such a simulated data set may be used in subsequent analyses.

The present study has some limitations. First, the number of items and the number of countries and the cultural diversity among them were limited. Second, the LC-BML model uses an optimization algorithm that may be quite time-consuming. This is especially true if the data set contains many observations and background variables, safeguards against local maxima are taken, and the numbers of segments are determined by optimizing an information criterion, so that the model parameters have to be optimized many times. Nevertheless, we have already successfully estimated the model on the PVQ (Schwartz et al. 2001) in Israel and 19 countries in Europe (Jowell et al. 2003), which comprises over 700,000 observations and has more predictor variables than the LOV data set; the results were roughly comparable to the results of the LOV data set. Third, the LC-BML model cannot be estimated using conventional, commercially available statistics packages, so that specialized programs had to be written. These programs have been written in the matrix programming language MATLAB and are available from the authors upon request.

Future research can build on our findings. The LC-BML model has been designed for data sets with substantive scales measured on rating scales as well as several background characteristics measured on nominal scales. Many data sets in empirical research, including marketing research, fit this format; future research could focus on applying the model to other data sets and further interpreting the model results. As the LC-BML model does not assume an ordering of the dependent variable Rating, it can, for example, also be used to analyze rating scale variables that include a 'don't know' category or have missing observations. By assigning a location to each rating category, our method may provide insight into the relative position of such categories. If a 'don't know' or 'missing' category is random, its location should be close to the origin in the plot; such a hypothesis may be tested using the confidence ellipses in the previous section. In short, we believe that the LC-BML model is a useful tool to provide more insight into how rating scale responses are affected by response styles, the content of the items, and background variables.

References

Anderson, Carolyn J. and Jeroen K. Vermunt (2000), "Log-multiplicative association models as latent variable models for nominal and/or ordinal data", *Sociological Methodology*, **30**, 81–122.

Anderson, John A. (1984), "Regression and Ordered Categorical Variables", *Journal of the Royal Statistical Society. Series B (Methodological)*, **46** (1), 1–30.

Andrews, Rick L. and Imran S. Currim (2003), "A Comparison of Segment Retention Criteria for Finite Mixture Logit Models", *Journal of Marketing Research*, **40** (2), 235–243.

Bachman, Jerald G. and Patrick M. O'Malley (1984), "Yea-saying, nay-saying, and going to extremes: Black-White differences in response style", *Public Opinion Quarterly*, **48** (2), 491–509.

Baumgartner, Hans and Jan Benedict E. M. Steenkamp (2001), "Response styles in marketing research: A cross-national investigation", *Journal of Marketing Research*, **38** (2), 143–156.

Bozdogan, Hamparsum (1987), "Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions", *Psychometrika*, **52** (3), 345–370.

Broen, W. E. Jr. and R. D. Wirt (1958), "Varieties of response sets", *Journal of Consulting Psychology*, **22** (3), 237–240.

Brunso, Karen, Joachim Scholderer, and Klaus G. Grunert (2004), "Closing the gap between values and behavior–a means-end theory of lifestyle", *Journal of Business Research*, **57** (6).

Chen, Chuansheng, Shin Y. Lee, and Harold W. Stevenson (1995), "Response style and cross-cultural comparisons of rating scales among East Asian and North American students", *Psychological Science*, **6** (3), 170–175.

Cheung, Gordon W. and Roger B. Rensvold (2000), "Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling", *Journal of Cross-Cultural Psychology*, **31** (2), 187–212.

Craig, C. Samuel and Susan P. Douglas (2000), *International marketing research*, Wiley, Chichester.

De Jong, Martijn G., Jan Benedict E. M. Steenkamp, Jean-Paul Fox, and Hans Baumgartner (2007), "Using Item Response Theory to Measure Extreme Response Style in Marketing Research: A Global Investigation", *Journal of Marketing Research*, **forthcoming**.

Demster, A. P., N. M. Laird, and D. B. Rubin (1977), "Maximum Likelihood from Incomplete Data via the EM-algorithm", *Journal of the Royal Statistical Society, Series B*, **39** (1), 1–38.

Fischer, Ronald (2004), "Standardization to Account for Cross-Cultural Response Bias: A Classification of Score Adjustment Procedures and Review of Research in JCCP", *Journal of Cross-Cultural Psychology*, **35** (3), 263–282.

Gifi, Albert (1990), *Nonlinear multivariate analysis*, Wiley, Chichester, England.

Greenleaf, Eric A. (1992a), "Improving rating scale measures by detecting and correcting bias components in some response styles", *Journal of Marketing Research*, **29** (2), 176–188.

— (1992b), "Measuring extreme response style", *Public Opinion Quarterly*, **56** (3), 328–351.

Groenen, Patrick J. F. and Alex J. Koning (2006), "A New Model for Visualizing Interactions in Analysis of Variance", in M. Greenacre and J. Blasius (eds.), *Multiple Correspondence Analysis and Related Methods*, Chapman & Hall, pp. 487–502.

Grunert, Klaus G., Susanne C. Grunert, and Sharon E. Beatty (1989), "Cross-Cultural Research on Consumer Values", *Marketing and Research Today*, **17** (1), 30–39.

Grunert, Susanne C. and Gerhard Scherhorn (1990), "Consumer Values in West Germany Underlying Dimensions and Cross-Cultural Comparison with North America", *Journal of Business Research*, **20** (2), 97–107.

Hamilton, David L. (1968), "Personality Attributes Associated with Extreme Response Style", *Psychological Bulletin*, **69** (3), 192–203.

Hofstede, Geert (2001), *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations*, Sage, Thousand Oaks, CA.

Hui, C. Harry and Harry C. Triandis (1989), "Effects of culture and response format on extreme response style", *Journal of Cross-Cultural Psychology*, **20** (3), 296–309.

Johnson, Timothy, Patrick Kulesa, Young I. Cho, and Sharon Shavitt (2005), "The Relation Between Culture and Response Styles: Evidence From 19 Countries", *Journal of Cross-Cultural Psychology*, **36** (2), 264–277.

Jowell, R. and the Central Co-ordinating Team (2003), "European Social Survey 2002/2003", Tech. rep., Centre for Comparative Social Surveys, City University, London.

Kahle, Lynn R. (1983), *Social values and social change: Adaptation to life in America*, Praeger, New York.

Kamakura, Wagner A. and Thomas P. Novak (1992), "Value-system segmentation: Exploring the meaning of LOV", *Journal of Consumer Research*, **19** (1), 119–132.

Kamakura, Wagner A. and Gary J. Russell (1989), "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure", *Journal of Marketing Research*, **26** (4), 379–390.

Krosnick, Jon A. and Duane F. Alwin (1988), "A test of the form-resistant correlation hypothesis: Ratings, rankings, and the measurement of values", *Public Opinion Quarterly*, **52** (4), 526–538.

Madrigal, Robert and Lynn R. Kahle (1994), "Predicting vacation activity preferences on the basis of value-system segmentation", *Journal of Travel Research*, **32** (3), 22–28.

Marin, Gerardo, Raymond J. Gamba, and Barbara V. Marin (1992), "Extreme response style and acquiescence among Hispanics: The role of acculturation and education", *Journal of Cross-Cultural Psychology*, **23** (4), 498–509.

Narayan, Sowmya and Jon A. Krosnick (1996), "Education moderates some response effects in attitude measurement", *Public Opinion Quarterly*, **60** (1), 58–88.

Paulhus, Delroy L. (1991), "Measurement and control of response bias", in John P. Robinson, Phillip R. Shaver, and Lawrence S. Wrightman (eds.), *Measures of Personality and Social Psychological Attitudes*, Academic Press, Inc., San Diego, CA, US, pp. 17–59.

Schwartz, Shalom H., Gila Melech, Arielle Lehmann, Steven Burgess, Mari Harris, and Vicki Owens (2001), "Extending the Cross-Cultural Validity of the Theory of Basic Human

Values with a Different Method of Measurement", *Journal of Cross-Cultural Psychology*, **32** (5), 519–542.

Smith, Peter B. (2004), "Acquiescent Response Bias as an Aspect of Cultural Communication Style", *Journal of Cross-Cultural Psychology*, **35** (1), 50–61.

Soutar, Geoffrey N., Richard Grainger, and Pamela Hedges (1999), "Australian and Japanese value stereotypes: A two country study", *Journal of International Business Studies*, **30** (1), 203–216.

Usunier, Jean-Claude and Julie A. Lee (2005), *Marketing Across Cultures*, Pearson Education Limited, Essex, England.

Van Herk, Hester (2000), *Equivalence in a cross-national context: methodological & empirical issues in marketing research*, PhD thesis, Tilburg University, Tilburg.

Van Herk, Hester, Ype H. Poortinga, and Theo M. M. Verhallen (2004), "Response Styles in Rating Scales: Evidence of Method Bias in Data From Six EU Countries", *Journal of Cross-Cultural Psychology*, **35** (3), 346–360.

Watson, Dorothy (1992), "Correcting for acquiescent response bias in the absence of a balanced scale", *Sociological Methods Research*, **21** (1), 52–88.

Wedel, Michel and Wagner A. Kamakura (2000), *Market Segmentation: Conceptual and Methodological Foundations*, 2nd edn., Kluwer, Dordrecht.

Wedel, Michel, Frenkel ter Hofstede, and Jan Benedict E. M. Steenkamp (1998), "Mixture Model Analysis of Complex Samples", *Journal of Classification*, **15** (2), 225–244.

Appendix A: Parameter Restrictions

Several parameter restrictions are required to ensure identification of the model parameters. Here, we give an overview of the restrictions that we impose.

First, we consider location restrictions, that is, restrictions that are necessary because adding a constant to a set of parameters would not change the estimated probabilities in the model. Location restrictions are needed for the parameters $a_{j|r}$, $\mathbf{F}$, $\mathbf{G}_k$, and $\mathbf{H}_s$. We impose sum-to-zero constraints for all these parameters per set, that is, we require that $\sum_{j=1}^{J} a_{j|r} = 0$, $\sum_{j=1}^{J} f_{jp} = 0$, $\sum_{t=1}^{T} h_{tp|s} = 0$, and that the elements of $\mathbf{g}_{kp}$ sum to zero for all values of k and p. As a consequence of these restrictions on $\mathbf{F}$, $\mathbf{G}_k$, and $\mathbf{H}_s$, the centroid of the points referring to a single variable (Rating, Item, or a background variable) is the origin in the plot.

Scale and rotation constraints are required for the parameters $\mathbf{F}$, $\mathbf{G}_k$, and $\mathbf{H}_s$. For notational convenience, let $\mathbf{B} = [\mathbf{B}_1|\ldots|\mathbf{B}_K]$, $\mathbf{C} = [\mathbf{C}_1|\ldots|\mathbf{C}_S]$, $\mathbf{G} = [\mathbf{G}_1'|\ldots|\mathbf{G}_K']'$, and $\mathbf{H} = [\mathbf{H}_1'|\ldots|\mathbf{H}_S']'$, where $\mathbf{C}_s$ contains the effects of the ratings on the items in item segment s. Then, the rank restrictions in Equation 2 can be written as

$$[\mathbf{B}|\mathbf{C}] = \mathbf{F}[\mathbf{G}'|\mathbf{H}']. \tag{8}$$

Without loss of generality, we can transform the parameter matrices as $\mathbf{F} = \mathbf{FT}$, $\mathbf{G} = (\mathbf{T}^{-1}\mathbf{G}')'$, and $\mathbf{H} = (\mathbf{T}^{-1}\mathbf{H}')'$ for any nonsingular $P \times P$ matrix $\mathbf{T}$, because

$$\mathbf{F}[\mathbf{G}'|\mathbf{H}'] = \mathbf{FTT}^{-1}[\mathbf{G}'|\mathbf{H}'].$$

This freedom of scaling also occurs in principal components analysis and correspondence analysis (Gifi 1990). We compute a constrained solution as follows. Let $[\mathbf{B}|\mathbf{C}]$ be obtained by some unconstrained $\mathbf{F}, \mathbf{G}$, and $\mathbf{H}$. In addition, let $\mathbf{P\Phi Q}' = [\mathbf{B}|\mathbf{C}]$ be a compact singular value decomposition with $\mathbf{P}$ and $\mathbf{Q}$ orthogonal rotation matrices with P columns (with $\mathbf{P}'\mathbf{P} = \mathbf{Q}'\mathbf{Q} = \mathbf{I}$) and $\mathbf{\Phi}$ a $P \times P$ diagonal matrix with positive monotonically decreasing values. As the rank of $[\mathbf{B}|\mathbf{C}]$ is not greater than P, matrices $\mathbf{P}$, $\mathbf{\Phi}$, and $\mathbf{Q}$ that meet the requirements above must exist and typically are unique up to a reflection per dimension. Here, we set $\mathbf{F} = \omega\mathbf{P}\mathbf{\Phi}^{1/2}$ and $[\mathbf{G}'|\mathbf{H}']' = 1/\omega\mathbf{Q}\mathbf{\Phi}^{1/2}$, where ω is a constant that determines the relative scaling of the points of the ratings compared to the points of the items and the background characteristics. The value of ω can be adapted without altering the general results. For graphical reasons, we choose ω so that the average squared Euclidean distance of the points to the origin is the same for both $\mathbf{F}$ and $[\mathbf{G}'|\mathbf{H}']'$, so that

$$\omega = \left(\frac{J}{ST + \sum_{k=1}^{K} m_k} \right)^{\frac{1}{4}}.$$

For the LOV data set, this choice amounts to setting $\omega \approx .65$. Furthermore, we simultaneously reflect the columns of $\mathbf{F}$, $\mathbf{G}$, and $\mathbf{H}$ in such a way that the first row of $\mathbf{F}$ only has positive values. The required identification constraints have now been obtained, while maintaining the equality in Equation 8, as

$$\mathbf{F}[\mathbf{G}'|\mathbf{H}'] = \omega\mathbf{P}\mathbf{\Phi}^{1/2} \left(\frac{1}{\omega}\mathbf{Q}\mathbf{\Phi}^{1/2} \right)' = \mathbf{P}\mathbf{\Phi}\mathbf{Q}' = [\mathbf{B}|\mathbf{C}].$$

Due to these parameter restrictions, the graphical representations in the Results section are scaled and rotated in such a way that the spread of the points in $\mathbf{F}$ and $[\mathbf{G}'|\mathbf{H}']'$ decreases with the dimension, and that the dimensions are orthogonal with respect to $\mathbf{F}$ and $[\mathbf{G}'|\mathbf{H}']'$. The point belonging to Rating 1 must have positive values on all dimensions.

Finally, the segment sizes are required to be nonnegative and nondecreasing, so that $u_1 \geq u_2 \geq \ldots \geq u_R \geq 0$ and $v_1 \geq v_2 \geq \ldots \geq v_S \geq 0$. In addition, they must sum to one, which implies that $\sum_{r=1}^{R} u_r = 1$ and $\sum_{s=1}^{S} v_s = 1$.

Table 6: Decomposition of the total number of degrees of freedom.

Source	Description	Degrees of freedom
$a_{j\|r}$	Attractiveness of ratings	$(J-1)R$
$\mathbf{F}$	Coordinates of Rating	$(J-1)P$
$\mathbf{G}_k$	Coordinates of background variables	$\sum_{k=1}^{K}(m_k-1)P$
$\mathbf{H}_s$	Coordinates of items	$S(T-1)P$
$\mathbf{u}$	Response style segment sizes	$R-1$
$\mathbf{v}$	Item segment sizes	$S-1$
Fixing rotation and scale freedom		$-P^2$

Now, the degrees of freedom of the model can be computed. Table 6 gives the number of degrees of freedom associated with each parameter set. The total number of degrees of freedom required by the LC-BML model equals

$$(J-1)(R+P)+\sum_{k=1}^{K}(m_k-1)P+S(T-1)P+R+S-2-P^2.$$

Appendix B: Optimization Algorithm

An Expectation-Maximization (EM) algorithm (Demster et al. 1977) is used to estimate the model parameters by maximizing the likelihood function. The EM algorithm starts with initial parameter estimates and then iteratively performs an E-step and a M-step, until convergence has been achieved. In the E-step, the posterior segment membership probabilities $\Pr(i \in \Xi_{r,s})$ are computed using Equation 6, given the current parameter estimates. In the M-step, the expected complete log-likelihood is maximized with respect to the parameter estimates, given the segment membership probabilities computed in the E-step. In our implementation, the M-step consists of 10 iterations of the BFGS quasi-Newton optimization routine in the MATLAB Optimization Toolbox (version 3.0.4), with analytically computed gradients. Convergence is considered to have been achieved if the change in log-likelihood between two consecutive EM iterations is smaller than 10^{-5}.

It is possible that the EM algorithm converges to parameter estimates that are only locally optimal. To solve this problem, the EM algorithm was rerun 10 times for every value of r and s with randomly chosen starting values, and the solution with the best likelihood value was retained.

Appendix C: Confidence Ellipses

The confidence ellipses in the Results section are based on maximum likelihood theory. Under regularity conditions, a maximum likelihood parameter estimator is normally distributed with covariance matrix equal to the information matrix. The information matrix can be estimated as the negative inverse of the matrix of second order partial derivatives of the likelihood function with respect to all model parameters, evaluated at the final parameter estimates. However, as the model parameters are not identified without the parameter constraints described in Appendix A, this matrix of second-order partial derivatives is not invertible. To circumvent this problem, we created an alternative model parametrization that does not require any parameter constraints for parameter identification and is equivalent to our original model.

The confidence intervals are constructed using the following procedure. First, the final parameter estimates of our original parametrization are transformed to the alternative one. Using these transformed parameter estimates, the matrix of second-order partial derivatives is calculated numerically. Then, $10,000$ simulated parameter vectors are drawn from a multivariate normal distribution with covariance matrix equal to the negative inverse of the matrix of second-order derivatives. For each simulated parameter vector, the associated parameter vector in the original parametrization is computed. Analysis of the simulated parameters in the original parametrization shows that the locations of the points in the graphical representation have approximately a joint bivariate normal distribution. Finally, the simulated locations of the points are used to construct 95% normal theory confidence ellipses.

Publications in the Report Series Research* in Management

ERIM Research Program: "Marketing"

2007

Marketing Communication Drivers of Adoption Timing of a New E-Service Among Existing Customers
Remco Prins and Peter C. Verhoef
ERS-2007-018-MKT
http://hdl.handle.net/1765/9405

Indirect Network Effects in New Product Growth
Stefan Stremersch, Gerard J. Tellis, Philip Hans Franses and Jeroen L.G. Binken
ERS-2007-019-MKT
http://hdl.handle.net/1765/9406

Demand-Driven Scheduling of Movies in a Multiplex
Jehoshua Eliashberg, Quintus Hegie, Jason Ho, Dennis Huisman, Steven J. Miller, Sanjeev Swami, Charles B. Weinberg and Berend Wierenga
ERS-2007-033-MKT
http://hdl.handle.net/1765/10069

Identifying Unknown Response Styles: A Latent-Class Bilinear Multinomial Logit Model
Joost van Rosmalen, Hester van Herk and Patrick J.F. Groenen
ERS-2007-045-MKT

---

* A complete overview of the ERIM Report Series Research in Management:
https://ep.eur.nl/handle/1765/1

ERIM Research Programs:
LIS Business Processes, Logistics and Information Systems
ORG Organizing for Performance
MKT Marketing
F&A Finance and Accounting
STR Strategy and Entrepreneurship