Ranking the ratings. A latent-class regression model to control for overall agreement in opinion research

Moors, G.B.D.

Published in:
International Journal of Public Opinion Research

Publication date:
2010

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Moors, G. B. D. (2010). Ranking the ratings. A latent-class regression model to control for overall agreement in opinion research. *International Journal of Public Opinion Research*, *22*(1), 93-119.

International Journal of Public Opinion Research Vol. 22 No. 1

doi:10.1093/ijpor/edp036 Advance Access publication 14 October 2009

RANKING THE RATINGS: A LATENT-CLASS REGRESSION MODEL TO CONTROL FOR OVERALL AGREEMENT IN OPINION RESEARCH

Guy Moors

Tilburg University, FSW-MTO, Room S110, PO Box 90153, 5000 LE Tilburg, the Netherlands

ABSTRACT

When rating questions are used to measure attitudes or values in survey research a researcher might want to control for the effect of overall agreement with the set of items that is rated. The need for controlling for overall agreement arises when the set of items refers to conceptual opposite perspectives, when balanced sets of items are used, or when a researcher is interested in relative preferences rather than overall agreement. In this paper, we introduce a method for filtering out overall agreement when a researcher's aim is to construct a latent class typology of respondents, that is, a latent-class ordinal regression model with random intercept. With this approach segments in the population are identified that differ in their relative preference of particular items over other items in the set. Examples are given on the concepts of locus of control, gender role attitudes and civil morality. The examples demonstrate that when an overall agreement is present in the data, the method is able to detect it, and at the same time allows identifying latent classes of respondents that differ in their relative preference of the items being rated.

The use of rating questions in measuring attitudes or values is still very popular. Rating questions involve a response format in which respondents are asked to indicate their level of agreement with particular issues. A typical format has been introduced by Likert (1932) who developed a response scale

Corresponding author: Guy Moors; e-mail: guy.moors@uvt.nl

This article was first submitted to *IJPOR* November 6, 2008. The final version was received July 6, 2009. This paper makes use of the CentERdata panel. The author gratefully acknowledges CentERdata and its director, Marcel Das, for including our short questionnaire.

with ordered categories ranging from 'completely disagree' to 'completely agree'. Although the number of response categories may vary, these rating questions have one thing in common: they aim at measuring the desirability or absolute value of a particular object. However, a researcher's interest might not be focused on absolute levels of agreement but rather on relative preferences of one issue over another. Take the following example: people may value a high income as well as opportunities for self-development in a job, but it is the preference of the first over the latter that defines the extrinsic versus intrinsic work motivation (Kohn & Schooler, 1983). From a substance point of view, then, it is useful to distinguish between an overall agreement response and the relative preferences of respondents. Next to this substance related justification for modelling overall agreement responses, a methodological motive has been provided. It has been argued that rating questions are susceptible for agreement bias or acquiescence (Billiet & McClendon, 2000). This type of response bias refers to the tendency among respondents to be more prone to choose one of the agreement response categories irrespective of the content of the issue (Paulhus, 1991). This can be easily demonstrated when a balanced set of items is developed to measure a particular latent construct. A balanced set of items includes positively and negatively worded items toward a particular attitude. The concept of ethnocentrism, for instance, can be measured by including items that refer to intolerance toward ethnic minorities as well as to items that reflect acceptance of other cultures (Billiet & McClendon, 2000). When respondents tend to agree with both types of issues, this can be regarded as agreement bias. Note that agreement tendency can only be interpreted as agreement bias when a balanced set of equivalent items is defined. With unbalanced scales, *e.g.*, only positively worded items, it is impossible to say whether the agreement tendency has a substantive meaning or whether it reflects a method bias.

In this article we demonstrate the usefulness of an approach that has been recently introduced in the context of consumer research to remove the effects of the overall response rating level (Magidson & Vermunt, 2006), *that is* a latent class regression model with random intercept. This approach is applicable when the aim of the researcher is to define latent segments or clusters in a population, and as such complements methods that have been developed for identifying latent dimensions or factors. We apply this approach in three examples: (1) questions referring to locus of control; (2) gender role attitudes; and (3) civil morality. The first example is chosen because of a conceptual need for distinguishing between intrinsic versus extrinsic locus of control. The second example includes balanced items and illustrates differential outcomes as far as agreement response bias is concerned. The third example demonstrates that it is possible to distinguish among clusters of respondents that differ in their relative ranking of morality issues after taking into account an overall agreement response pattern.

The article is organized as follows. First, we elaborate on the method that was developed for removing the effect of overall response rating in estimating taste preferences of consumers and argue how this perspective can be generalized to attitude research. Secondly, we define the context or situation in which the regular survey researcher might consider using this approach. And finally, we apply the method with examples from these particular situations to demonstrate the usefulness of the approach.

RANKING THE RATINGS: INTRODUCING A METHOD TO REMOVE AN OVERALL RESPONSE RATING LEVEL IN CONSUMER RESEARCH AND—BY EXTENSION—IN SURVEY RESEARCH

Which type of cornflakes is preferred by which segment of the population? Are there particular design aspects of a car that influence a young woman's choice in buying that car? We all can imagine the kind of questions that are relevant in consumer research. In this type of research there is a primary need to understand differences in taste preferences of consumers. This is important when new products need to be developed and/or when advertising aims at a particular consumer segment. The predominant method of data collection (Magidson & Vermunt, 2006), however, asks a test panel to rate each of the attributes of a product as such. The popularity of rating scales lies in their convenience (van Herk & van de Velden, 2007; see also Munson & McIntyre, 1979)–rating scales are easy to administer, can be completed in a short period of time, and are usually not difficult to understand for respondents. Furthermore, they allow for use of parametric statistical methods. Other formats such as preference ranking and pair-wise comparison, by contrast, are often considered to be too demanding for respondents (Klien, Dülmer, Ohr, Quandt & Rosar, 2004); although it has been argued that the use of discrete choice experiments to elicit preferences may decrease response burden (Ryan, Bate, Eastmond & Ludbrook, 2001). Nevertheless, rating question remains the predominant method.

The problem with rating data in consumer research, however, is that an overall liking tends to dominate the results rather than that one captures differences in preference between the presented products (Magidson & Vermunt, 2006). This overall liking reflects a respondents' response tendency that should be taken into account when the principal aim of a researcher is to get information on the preference of a product relative to other products. Note that when we are referring to the overall liking as a response tendency we do not imply this to be a response bias. Such a tendency is only a response bias if it is independent from the true content of what the researcher aims at measuring. There are many cases, however, in which an overall liking (or disliking) makes perfectly sense. Personally, for instance, I don't like cornflakes, but as a test person I can rate different cornflakes and even make a distinction in taste between them.

To eliminate the overall response level effect in rating data it has been suggested that a within-case 'centering' of the data might be useful (Cattell, 1944; Cunningham, Cunningham & Green, 1977; Magidson & Vermunt, 2006). This within-case centering involves the calculation of (a) a mean level of liking for all objects (items) that are rated and (b) subtracting this value from the observed value for each object (item). As such it measures the deviance from the personal average liking of the objects or items that are rated. Positive values indicate higher than average liking, negative values indicate less than average liking. It has been pointed out (Dunlap & Cornwell, 1994; Cheung & Chan, 2002; Cheung, 2006) that from a statistical point of view this procedure causes particular problems that relate to the ipsative nature of within-case centering data. The measurement of one object (item) is no longer independent from the measurement of other objects (items) in the set since the sum of all values of the within-case centering variables is a constant (=0). With k items in a set only k-1 within-case centering variables are needed for full information on the relative rating of all k items. If, however, a researcher needs information on relative preferences rather than overall liking a model-based alternative to within-case centering of rating data should be considered. As is demonstrated by Magidson & Vermunt (2006) a latent-class ordinal regression model with random intercept provides such a model-based alternative. Furthermore, this approach has the additional benefit of maintaining the ordinal discrete metric of the observed rating data. After all, when an individual's overall mean is subtracted from the discrete ratings of the items the original discrete distribution changes in a kind of continuous scale with a complicated distribution (Magidson & Vermunt, 2006; Popper, Kroll & Magidson, 2004). Furthermore, Popper, Kroll & Magidson (2004) have demonstrated that the LC ordinal regression model with random intercept outperforms alternative approaches that also aim at distinguishing between an overall liking dimension and a relative preference dimension.

We use Latent GOLD 4.5 to analyze the models presented in this article. An example data layout for the latent class ordinal regression model is presented in figure 1. The layout reflects a multilevel data structure in which each rated object or item (=lowest level) defines a separate record for each person (Respondent ID = highest level). A nominal variable (Item ID) identifies which item is rated and an ordinal response variable (Rating) identifies the rating that is given to a particular item. Individual level covariates, such as gender, age or education, may also be included.

Let Y_{ij} denote the rating of respondent i of item j, with $i = 1, 2, \ldots, n$ (n = number of respondents); and $j = 1, 2, \ldots, m$ (m = number of items). The rating (Y_{ij}) takes on discrete values denoted by k with $k = 1, 2, \ldots,$

FIGURE 1 **Example data layout for the LC ordinal regression models**

Respondent ID	Item ID	Rating	Gender	Age
1	do anything	4	male	32
1	responsible for failure	2	male	32
1	bad luck	3	male	32
1	good fortune	3	male	32
2	do anything	5	female	41
2	responsible for failure	5	female	41
2	bad luck	4	female	41
2	good fortune	3	female	41

r (r = number of response categories). Given that the rating is a discrete ordinal response variable we define an adjacent-category logit model (Agresti, 2002) as follows:

$$\log\left[\frac{P(Y_{ij} = k|x)}{P(Y_{ij} = k-1|x)}\right] = \alpha_{ik} + \beta_{xj}$$

$$\text{with } \alpha_{ik} = \alpha_k + \lambda F_i$$

This is the specification of a regression model for the logit associated with a given rating m instead of $m-1$ for item j conditional on membership of latent class x, for $x = 1, 2, \ldots, t$ (t = number of latent classes). The intercept, α_{ik} is allowed to vary across individuals and is defined as a function of the expected value of the intercept ($=\alpha_k$) and a normally distributed continuous factor (F_i) which has a mean equal to 0 and a variance equal to 1, and where λ is the factor loading. Given that the variance of the intercept (α_{ik}) equals λ^2, both the expectation and the square root of the variance are model parameters. The β_{xj} parameter in the equation refers to the effect of the jth item for latent class x. By using effect coding for parameter identification, the β_{xj} sum to zero over all items so that positive values for β_{xj} indicate that a particular latent class x likes an item more than average and negative values indicate that the item is less preferred than average. It is clear that in this model the random intercept accounts for the overall agreement tendency, whereas the β_{xj} parameters indicate the relative preference of an item compared to the average liking of all items.

WHEN IS CORRECTING FOR OVERALL AGREEMENT IN PUBLIC OPINION RESEARCH USEFUL?

Public opinion research is rarely about taste preferences, although one could argue that a similarity can be drawn between rating products and expressing an opinion by rating attitude questions. As indicated in the introduction there are different situations in which a public opinion researcher might wish to correct for overall agreement tendencies in attitude or values measurement. These reasons may be of a more conceptual nature, or they may relate to methodological concerns.

First, social scientists often develop sets of items to measure attitudes and values that are assumed to measure opposite opinions. Well-known examples in the literature are: (a) the concepts of 'left' versus 'right' or 'materialism' versus 'postmaterialism' in political science (Inglehart, 1990); or (b) 'internal' versus 'external' motivation in social psychology (Rotter, 1954); and (c) 'self-directedness' versus 'external benefits of a job' in sociology (Kohn & Schooler, 1983). We admit that it has been debated whether these concepts are truly bipolar or rather distinct concepts that do not necessarily oppose one another (Alwin & Krosnick, 1988). At the same time, we subscribe to the principle that when a researcher defines a concept in a particular way; the operationalization should be consistent with the concept. This is exactly the argument raised by Inglehart (1990) among others, who claims that in politics—as in life—it is about making choices between alternatives and, by consequence, attitude research should focus on these choices people make. Asking ranking questions, in this case, would be the obvious choice of a researcher. After all, a ranking assignment involves respondents indicating their first, second, *etc.* choice among a set of alternatives. There are, however, instances in which ranking cannot be used. In secondary data analysis, for instance, a researcher is left without a choice. Also, the complexity of a ranking assignment becomes problematic when the list of options is fairly long. Suggested alternatives for long lists of items such as pair-wise comparisons or using triads, also cause burden upon respondents since it increases survey time. In these cases a researcher can use the suggested method to control for overall agreement and develop a measurement model that reflects relative preferences. Since the method involves defining segments in the population that differ in their preference structure, the method is applicable if the researcher's main interest is in classifying respondents that differ in their preference profiles. Hence, the approach is conceptually similar to cluster analysis in that it aims at classifying similar objects into groups of which the number of groups as well as their forms are unknown (Kaufman & Rousseeuw, 1990). In this article we apply the method on four items referring to the concept of internal-external locus of control (Rotter, 1954; Mirowsky & Ross, 1990).

Secondly, an overall agreement tendency translates into agreement bias when a balanced set of agreement scales is used to measure a particular construct. In this case respondents reveal a tendency to opt for the agreement categories of a response scale regardless whether the attitude questions are positively or negatively worded. In this article a balanced set of gender role attitudes is analyzed. The methodological concern that inspired researchers to develop balanced sets of questions is the knowledge that (some) respondents tend to agree (*yes-saying*) in answering survey questions irrespective of the content of the items. This phenomenon especially occurs when large sets of items with dichotomous response alternatives (yes/no; agree/disagree; true/false) are used (Krosnick, Judd & Wittenbrink, 2005). However, skewness toward the agreement categories of Likert type of questions is also indicative of an agreement response style (Billiet & McClendon, 2000). To correct for agreement bias in the latter case, Billiet & McClendon (2000) have developed an approach in the context of structural equation modeling that looks solid. Their approach, however, involves a confirmatory factor analysis and is, by consequence, applicable when a researcher focuses on latent factors or dimensions. The method that we have described in the previous section is different in the sense that it identifies latent classes, which is conceptually more similar to identifying clusters. Hence, a researcher who wants to identify a gender role typology (Eid, Langeheine & Diener, 2003) may consider adopting the approach suggested in this article. Finally, our third application of the model on a set of issues reflecting different aspects of civil morality (Halman, 1997) provides a nice example on how these issues are rank ordered within different segments of the population.

In summary, we suggest that a public opinion researcher should seriously consider the use of the approach elaborated in this article when he or she:

(a) aims at classifying respondents in latent classes or clusters in terms of their relative preference of items compared to other items in a set of questions,
(b) is using or has to use rating scales, and
(c) needs to control for overall agreement.

SAMPLING AND DATA

Data used in this research come from a short questionnaire that was designed and administered in the context of the Dutch CentERpanel web-survey that was established in 1991. This panel started as a random sample of over 2000 households in the Netherlands representative of the Dutch population. Panel members, aged 16 years or older, complete a questionnaire on the internet from their home on a weekly basis. Households without internet access are supplied with a set-top box with which questionnaires can be completed using a television screen as a monitor. Since sampling strategy is random, this internet panel is not composed of self-selected members, but rather a sample of respondents that represents the full population of Dutch households.

We lack information to calculate different response rates at different stages of this panel research in the period from 1991 to the date of our measurement in December 2005. Important to know, however, is that the organizers of this panel offered us the facility to administer a short questionnaire to a subsample of all panel members in the mailing list of December 2005. A random subsample of 1468 respondents was contacted of which 1051 respondents filled in the questionnaire. Hence, the refusal rate of all contacted persons was 29 percent.

The analyses reported in this article are conducted with Latent GOLD 4.5. This program also allows specifying information on the sampling design and accordingly corrects standard errors and Wald statistics (Vermunt & Magidson, 2005). In our models standard errors and Wald statistics are corrected for the clustering of individuals within household.

The short questionnaire included two sets of four questions referring respectively to 'gender roles', and 'locus of control'. Respondents were asked to rate their agreement on a six-point scale ranging from 1 = 'completely disagree', to 6 = 'completely agree'. A third set of six questions referred to issues of 'civil morality' and respondents needed to indicate on a scale from 1 = 'never' to 10 'always' to what extent they rated particular behaviours as justified. A 'no opinion' option was not presented on the screen because this was the regular practice in this internet panel. The use of short questionnaires is standard practice in this (and other) internet panel research. Furthermore, it is not uncommon in public opinion research to use short scales, especially not when information is collected on a broad range of topics. The method proposed in this article, however, can be used with longer sets of questions as well.

The questionnaire was administered in Dutch and the selected items were adapted from regular surveys as indicated in the first table. Table 1 presents an overview of all items per concept and their associated label that will be used in the remainder of this article.

Before presenting the results from the LC regression models with random intercept we need to present some details of our analyses. The first step in any latent class analysis usually involves estimating latent class models with 1, $2, \ldots, n$ number of latent classes and comparing the fit of each one to the data. Model fit is provided by its log likelihood (LL) and associated Bayesian Information Criterion (BIC) value, which is a parsimony index. The lower this latter value, the better the model fit. The principal reason to use information criterion in model selection is that it allows for comparisons of non-nested models, which is the case when latent classes are added. We will present these model statistics for LC regression models with and without random intercept. The differences between these two models is that the latter is a traditional LC regression model in which latent classes differ with respect to both the

TABLE 1 **Overview of items (items ordered as in the questionnaire)**

(a) Gender role attitudes (adapted from the European Values Survey)	*Label*	
a1. A working mother can establish just as warm and secure a relationship with her children as a mother who does not work.	Working mother	(−)
a2. A pre-school child is likely to suffer if his or her mother works.	Pre-school child	(+)
a3. A job is alright but what most women really want is a home and children.	Family orientation	(+)
a4. There is more in life than family and children, what a woman also needs is a job that gives her satisfaction.	Job orientation	(−)

Note: (+) refers to gender role stereotyping (inequality); (−) refers to emancipated attitudes (equality); response categories from 1 = completely disagree to 6 = completely agree.

(b) Locus of control (adapted from Mirowsky & Ross, 1990)	*Label*	
b1. I can do just about anything I really set my mind to.	Insuccess	(I S)
b2. I am responsible for my failures.	Infail	(I F)
b3. Most of my problems are due to bad breaks.	Fatfail	(E F)
b4. The really good things that happen to me are mostly luck.	Fatsuccess	(E S)

Note: Mirowsky & Ross (1990) assign a 2-folded meaning to each issue, i.e. 'I' = internal locus; 'E' = external locus; 'S' = success; and 'F' = failure; response categories from 1 = completely disagree to 6 = completely agree.

(c) Civil morality (adapted from the European Values Survey)	*Label*	
c1. Cheating on tax if you have the chance	Cheating on tax	(P E)
c2. Claiming state benefits which you are not entitled to	Claiming state benefits	(P E)
c3. Paying cash for services to avoid taxes	Paying cast to avoid taxes	(P E)
c4. Abortion		(L D)
c5. Euthanasia		(L D)
c6. Homosexuality		(L D)

Note: (PE) = 'personal enrichment'; (LD) = 'interference in life and death of people'; response categories from 1 = never justified to 10 = always justified.

intercept and regression coefficients, whereas in the former model individual-level variability in the intercept is captured by the random intercept. The comparison of these models with and without random intercept allows evaluating the significance of including the random intercept in our models. Additionally we will compare the LC standard pseudo R^2 statistics of the models. Details on the selected model will be presented for each example. In the second step of the analysis, we present the results from the selected model. Each of the three example models includes a set of covariates, *i.e.* gender, age (in categories), education, living arrangement and job status. As such our models fit within a structural equation modeling (SEM) framework, which is of course a regular practice in public opinion research. Finally, we cross validate the results of the selected LC regression models with random intercept with the results from a regular latent class approach in which a "counting for agreement variable" was included as a control variable. This latter variable is operationalized as the total number of agreement responses after dichotomizing items contrasting the disagreement response options (=0) with the agreement options (=1) (Billiet & McClendon, 2000; see also: Ray, 1979). If the random intercept adequately captures overall agreement, than the results of the two models should be similar. The major difference, however, is that the model-based estimate of overall agreement is statistically more solid since it is a probabilistic model. This implies that for each respondent a probability of having given an overall agreement response is estimated, whereas the count variable may be biased by measurement error, especially when only a few items are included in this count index. The detailed results of the "counting for agreement" adjustment models are not presented since in each example the similarity in results was obtained. The major differences between the random intercept model and a model controlling for the count variable were that standard errors of the measurement model were generally smaller in the intercept model and that the estimated proportion of respondents in each latent class slightly shifted. As a summary result we have listed the correlation between the random intercept scores at the individual level with the "counting for agreement" variable.

DEMONSTRATING THE USEFULNESS OF A LATENT-CLASS ORDINAL REGRESSION MODEL WITH RANDOM INTERCEPT

For each example the results in the tables are presented in three parts. First, we compare the model selection and fit statistics of consecutive models with and without random intercept. Second, the latent class estimates are presented showing the expected value of the intercept (α_k), the random intercept loading (λ) and the effect sizes of items for each latent class (β_{xj}). The third part of the table includes the structural part of the model in which the effects of covariates on latent class membership are presented. Significance at the variable level is indicated by the Wald statistic, and standard errors allow to evaluate the significance of parameters (parameter $> 1.96 \times$ SE). The question on 'what difference does it make' is answered at the end by comparing results with outcomes from the standard latent class regression analyses.

Example 1: Locus of Control

The first example includes a selection of four items referring to locus of control as presented in a study by Mirowsky & Ross (1990). In their work Mirowsky & Ross link these issues to the concepts of control and defence. They argue that locus of control gets a different meaning whether a respondent's defence mechanism is self-defensive or self-blaming. A self-defensive mechanism occurs when a respondent claims responsibility for success or when he/she denies responsibility for failure. Self-blaming refers to a situation in which a respondent claims responsibility for failure or denies responsibility for success. The interesting thing about Mirowsky & Ross' arguments from the perspective of our article is that they provide a theoretical rationale for applying latent class analysis. After all, they define a theoretical typology and the aim of a latent class analysis is identifying segments in populations and as such defining an empirical typology.

In both the standard latent-class regression model and the random intercept model the two class model reveals the lowest BIC value. The two class random intercept model performs better: overall it has the lowest observed BIC and the pseudo R^2 is substantially higher (=.67) than in the model without random intercept (=.45).

Examining the latent class part of the selected model we find confirmation of a strong positive random intercept effect. This is consistent with the argument that overall agreement tends to dominate the picture, as is confirmed by the correlation of this random intercept with the count variable indicating the number of agreement scores on this set of items ($r = .791$). The positive random intercept effect implies that respondents indicated that all four issues are influential factors in the lives of people. The latent classes identify the relative preferences of particular issues compared to others controlling for the average agreement (*i.e.* random intercept) with the four issues. The two latent classes claim responsibility for both successes (insuccess) and failures (infail). Mirowsky & Ross (1990) made similar observations. The main difference between the first and second class is that the first latent class only marginally makes that distinction between claiming responsibility and denying responsibility, whereas in the second latent class this difference is more clearly articulated.

TABLE 2 **Locus of control: a LC-regression model with random intercept**

2.1 Model selection and fit statistics

	Loglikelihood (LL)	*BIC (LL)*	*Standard pseudo R^2*
1-Class regression	−5869.3	11794.2	0.32
2-Class regression	−5788.3	11771.4	0.45
3-Class regression	−5765.6	11865.2	0.45
1-Class + random intercept	−5812.1	11686.8	0.48
2-Class + random intercept	−5663.2	11528.2	0.67
3-Class + random intercept	−5627.0	11594.9	0.71

2.2 Latent class part of the model (measurement of the model)

Intercept	*α (SE)*	*Wald*	*P-value*
Completely disagree	−3.392 (0.178)	580.097	.000
Disagree	0.308 (0.065)		
Rather disagree	1.625 (0.090)		
Rather agree	2.085 (0.100)		
Agree	1.323 (0.069)		
Completely agree	−1.948 (0.137)		

Random intercept (agreement response pattern)	*λ (SE)*	*Wald*	*P-value*
	0.697 (0.059)	139.469	.000

Latent classes	*Class 1*	*Class 2*		
Items	*β (SE)*	*β (SE)*	*Wald*	*P-value*
Insuccess	0.195 (0.058)	1.534 (0.182)	306.931	.000
Infail	0.954 (0.073)	2.538 (0.177)		
Fatfail	−0.616 (0.053)	−1.988 (0.197)		
Fatsuccess	−0.533 (0.075)	−2.084 (0.166)		
Class size	*Proportion (SE)*	*Proportion (SE)*		
	0.709 (0.049)	0.291 (0.049)		
LC labels	Class 1 = high on claiming versus low on denying responsibility–moderate ranking Class 2 = high on claiming versus low on denying responsibility–strong ranking			

TABLE 2 **Continued**

2.3 Structural part of the model (predicting class membership)

Covariates / *Latent classes*	*β (SE)* *Class 1*	*β (SE)* *Class 2*	*Wald*	*P-value*
Gender				
Men	−0.073 (0.048)	0.073 (0.048)	2.290	.130
Women	0.073 (0.048)	−0.073 (0.048)		
Age (in categories), years				
15–24	0.117 (0.181)	−0.117 (0.181)	1.569	.900
25–34	−0.018 (0.123)	0.018 (0.123)		
35–44	0.028 (0.113)	-0.028 (0.113)		
45–54	−0.107 (0.102)	0.107 (0.102)		
55–64	−0.054 (0.117)	0.054 (0.117)		
≥65	0.034 (0.153)	−0.034 (0.153)		
Educational level				
Primary education	0.572 (0.273)	−0.572 (0.273)	23.930	.000
Lower secondary education (vmbo)	0.182 (0.110)	−0.182 (0.110)		
Higher secondary education (havo/vwo)	−0.156 (0.129)	0.156 (0.129)		
Intermediate vocational training (mbo)	0.113 (0.122)	−0.113 (0.122)		
Higher vocational training (hbo)	−0.260 (0.099)	0.260 (0.099)		
University (wo)	−0.452 (0.135)	0.452 (0.135)		
Living arrangement				
Alone	−0.046 (0.139)	0.046 (0.139)	0.874	.930
With a partner, no children	−0.052 (0.123)	0.052 (0.123)		
With a partner and children	−0.098 (0.119)	0.098 (0.119)		
Alone with children	−0.069 (0.220)	0.069 (0.220)		
Other	0.265 (0.316)	−0.265 (0.316)		
Doing paid work				
No	0.062 (0.067)	−0.062 (0.067)	0.842	.360
Yes	−0.062 (0.067)	0.062 (0.067)		

Only education is significantly related to the latent class typology of locus of control (P-value of the Wald-statistic $<.00$). Education is a complex categorical variable distinguishing between primary (first category), secondary (second and third category) and post-secondary levels. Within this post-secondary level the three categories are also listed by increasing level of education. At the same time the type of education (general versus vocational training) also defines the classification. For example, depending on the track a student has followed at the secondary level he or she is more likely to progress to a particular type of higher post-secondary level. Higher secondary education—which is a general training—prepares for university. As such the classification of educational categories should be thought of as both a vertical (up-down levels) and horizontal (differences in type of education) stratification. With this in mind the general picture that emerges from the analysis is that the likelihood of being in the second latent class, which more strongly claims responsibility for both successes and failures, increases by level and by type of training (general versus vocational). Among the first three categories of education the likelihood of being classified in the second latent class increases. This likelihood also increases with increasing levels within the post-secondary levels. The fact that respondents with a higher secondary training have a somewhat higher likelihood of being in the second latent class compared to the intermediate vocational training level indicates that a distinction in vocational-general education also plays a role.

Example 2: Gender Role Attitudes

The second example is chosen from a longstanding theme in sociology, *i.e.* gender role attitudes (Thornton & Young-DeMarco, 2001). The set of questions includes two items that reflect gender stereotyping and two items that indicate gender equality. These two sets of items are regarded as conceptually balanced, and hence as items located at the poles of a continuum. Two items reflect opposite views on the work-family balance in the life of women ('job' vs. 'family orientation') and two items discuss the impact of having paid work while being a mother ('working mother' versus 'pre-school child').

Comparing the model selection criteria we observe a similar result as in the previous example. BIC values indicate that in both types of models, with and without random intercept, the same number of latent classes is selected. In this example a three class model is preferred. Again the random intercept model performs better, although only marginally so: R^2 only increases with .05. This should not come as a surprise since the random intercept has a different meaning compared to the previous example. In the locus of control example overall agreement with the four issues can be fully legitimate from a content point of view. As far as the gender role attitude questions are concerned an overall agreement tendency much more reflects a response bias.

TABLE 3 Gender role attitudes: LC-regression model with random intercept

3.1 Model selection and fit statistics

	Loglikelihood (LL)	*BIC (LL)*	*Standard pseudo* R^2
1-Class regression	−6965.7	13987.1	0.12
2-Class regression	−6514.5	13223.8	0.50
3-Class regression	−6418.5	13170.9	0.56
4-Class regression	−6354.6	13182.3	0.57
1-Class + random intercept	−6965.7	13994.0	0.12
2-Class + random intercept	−6514.0	13229.7	0.51
3-Class + random intercept	−6411.8	13164.5	0.61
4-Class + random intercept	−6343.4	13166.9	0.64

3.2 Latent class part of the model (measurement of the model)

Intercept	α *(SE)*	*Wald*	*P-value*
Completely disagree	−2.099 (0.110)	472.315	.000
Disagree	−0.019 (0.046)		
Rather disagree	0.710 (0.055)		
Rather agree	1.240 (0.061)		
Agree	0.807 (0.052)		
Completely agree	−0.638 (0.066)		

Random intercept (agreement response pattern)	λ *(SE)*	*Wald*	*P-value*
	0.226 (0.049)	21.358	.000

Latent classes	*Class 1*	*Class 2*	*Class 3*		
Items	β *(SE)*	β *(SE)*	β *(SE)*	*Wald*	*P-value*
Working mother	0.177 (0.170)	2.045 (0.167)	−1.888 (0.655)	526.558	.000
Pre-school child	0.023 (0.159)	−1.902 (0.179)	1.867 (0.536)		
Family priority	−0.249 (0.073)	−1.161 (0.102)	0.555 (0.301)		
Job importance	0.049 (0.062)	1.019 (0.107)	−0.534 (0.202)		
Class size	*Proportion (SE)*	*Proportion (SE)*	*Proportion (SE)*		
	0.559 (0.027)	0.347 (0.045)	0.094 (0.047)		
LC labels	Class 1 = non-differentiators Class 2 = egalitarian gender roles Class 3 = gender stereotyping				

(continued)

TABLE 3 **Continued**

3.3 Structural part of the model (predicting class membership)

Covariates *Latent classes*	*β (SE)* *Class 1*	*β (SE)* *Class 2*	*β (SE)* *Class 3*	*Wald*	*P-value*
Gender					
Men	0.071 (0.070)	−0.366 (0.069)	0.295 (0.107)	32.131	.000
Women	−0.071 (0.070)	0.366 (0.069)	−0.295 (0.107)		
Age (in categories), years					
15–24	0.364 (0.250)	−0.051 (0.313)	−0.313 (0.468)	9.939	.450
25–34	0.184 (0.208)	−0.141 (0.181)	−0.042 (0.318)		
35–44	0.043 (0.168)	−0.037 (0.165)	−0.006 (0.269)		
45–54	−0.125 (0.141)	−0.030 (0.142)	0.155 (0.225)		
55–64	−0.218 (0.160)	0.195 (0.159)	0.024 (0.226)		
≥65	−0.246 (0.185)	0.064 (0.225)	0.183 (0.291)		
Educational level					
Primary education	−0.039 (0.234)	−0.110 (0.239)	0.149 (0.345)	35.645	.000
Lower secondary education (vmbo)	0.251 (0.132)	−0.544 (0.139)	0.293 (0.193)		
Higher secondary education (havo/vwo)	−0.128 (0.213)	0.209 (0.162)	−0.081 (0.306)		
Intermediate vocational training (mbo)	0.171 (0.145)	−0.220 (0.146)	0.049 (0.219)		
Higher vocational training (hbo)	0.033 (0.129)	0.223 (0.124)	−0.256 (0.188)		
University (wo)	−0.288 (0.216)	0.441 (0.177)	−0.154 (0.302)		
Living arrangement					
Alone	−0.076 (0.249)	0.051 (0.245)	0.025 (0.420)	18.976	.015
With a partner, no children	−0.558 (0.210)	−0.220 (0.249)	0.779 (0.380)		
With a partner and children	−0.594 (0.223)	−0.128 (0.261)	0.722 (0.366)		
Alone with children	1.115 (0.417)	0.558 (0.506)	−1.673 (0.773)		
Other	0.113 (0.621)	−0.261 (0.639)	0.148 (1.061)		
Paid job					
No	0.198 (0.095)	−0.317 (0.101)	0.119 (0.143)	15.131	.001
Yes	−0.198 (0.095)	0.317 (0.101)	−0.119 (0.143)		

In general one would expect that when respondents answer positively and negatively worded questions consistent with the true content of the concept, then the random intercept would be 0 or insignificant. In the case of gender role attitudes, however, there is a significant positive random intercept indicating an agreement tendency. This is confirmed in the 'counting for agreement' model. Again the random intercept is strongly correlated with the counting for agreement variable ($r = .770$).

The latent class profile is consistent with a conceptual framework that distinguishes between an egalitarian and a gender stereotyping orientation, respectively the second and third latent class. The second egalitarian latent class shows relatively higher than average rating of 'job importance' and higher agreement with the issue that a 'working mother' can establish a good relationship with her children as well. At the same time the average rating of 'a preschool child' suffering from having a working mother and 'family priority' is less than average. The third gender stereotyping latent class defines the counterpart. The first and largest latent class, however, hardly makes any difference between the gender issues, except for the issue of 'family priority', which is significantly less preferred. This means that this class qualifies as a category of non-differentiators. Non-differentiators are respondents who give (nearly) identical scores to each item. This interpretation is confirmed by the finding that the latent class probabilities of belonging to this first latent class strongly correlate with an index calculated as the standard deviation of the four scores given by an individual ($r = -.82$).

Not surprisingly a gender cleavage is observed in the likelihood of being classified in the second and third latent class with women being more egalitarian minded (second class) and men being more conservative (third class).

Overall education is significantly related to the latent class typology, but the differences between educational categories are only significant as far as the second egalitarian class is concerned. University students are clearly more egalitarian in gender roles than other categories, whereas the lower secondary educational level is the least likely to be classified in this second latent class. Again the level-type combination in education proves to be important.

In contrast to education, living arrangements are related to the probability of belonging to the first 'neutral' and third 'gender stereotyping' latent classes. Respondents living with a partner (with or without children) contrast with single parents on these issues. The former are more likely to use 'gender stereotyping' in their attitudes, whereas single parents are much less likely. The latter, by contrast, are more likely to be classified in the first 'neutral' category, whereas respondents living with a partner are not.

Having a paid job is linked to classification in the first 'neutral' category and the second 'egalitarian' latent class, with respectively a negative and positive estimated likelihood.

TABLE 4 **Civil morality: LC-regression model with random intercept**

4.1 Model selection and fit statistics

	Loglikelihood (LL)	*BIC (LL)*	*Standard Pseudo R^2*
1-Class regression	−12137.2	24371.8	0.35
2-Class regression	−11768.9	23788.4	0.52
3-Class regression	−11683.9	23771.3	0.54
4-Class regression	−11651.5	23859.6	0.55
1-Class + random intercept	−11750.5	23605.4	0.62
2-Class + random intercept	−11292.6	22842.6	0.77
3-Class + random intercept	−11190.6	22791.7	0.80
4-Class + random intercept	−11144.0	22851.6	0.82

4.2 Latent class part of the model (measurement of the model)

Intercept	*α (SE)*	*Wald*	*P-value*		
Never 1	0.940 (0.108)	504.337	.000		
2	0.429 (0.103)				
3	0.596 (0.089)				
4	0.350 (0.083)				
5	1.122 (0.063)				
6	0.470 (0.067)				
7	0.164 (0.079)				
8	−0.255 (0.089)				
9	−1.672 (0.116)				
Always 10	−2.144 (0.153)				
Random intercept (agreement response pattern)	*λ (SE)*	*Wald*	*P-value*		
	0.474 (0.019)	599.914	0.000		
Latent classes	*Class 1*	*Class 2*	*Class 3*		
Items	*β (SE)*	*β (SE)*	*β (SE)*	*Wald*	*P-value*
Claiming state benefits not entitled to	−1.271 (0.079)	−0.587 (0.056)	−4.289 (0.910)	908.589	.000
Cheating on tax	−0.337 (0.049)	0.114 (0.039)	−0.756 (0.088)		
Paying cash for services to avoid taxes	−0.011 (0.043)	0.514 (0.050)	−0.212 (0.069)		
Abortion	0.074 (0.036)	−0.182 (0.037)	0.991 (0.182)		
Euthanasia	0.476 (0.034)	0.116 (0.041)	1.226 (0.163)		
Homosexuality	1.069 (0.098)	0.024 (0.065)	3.040 (0.571)		
Class size	*Proportion (SE)*	*Proportion (SE)*	*Proportion (SE)*		
	0.545 (0.026)	0.265 (0.037)	0.191 (0.040)		
LC labels	Class 1 = stepwise justification of actions (from disapproving personal enrichment to approving personal interference in life) Class 2 = non-differentiators (only differentiating issues of 'state benefits' (disapproval) and 'justifying paying cash' (approval) Class 3 = contrasting acceptance of personal enrichment (lowest) to interference in personal life (highest)				

Table 4 **Continued**

4.3 Structural part of the model (predicting class membership)

Covariates *Latent classes*	*β (SE)* *Class 1*	*β (SE)* *Class 2*	*β (SE)* *Class 3*	*Wald*	*P-value*
Gender					
Men	−0.134 (0.062)	0.396 (0.073)	−0.261 (0.080)	29.824	.000
Women	0.134 (0.062)	−0.396 (0.073)	0.261 (0.080)		
Age (in categories), years					
15–24	−0.050 (0.195)	−0.110 (0.212)	0.160 (0.248)	19.269	.037
25–34	−0.101 (0.151)	0.212 (0.158)	−0.111 (0.166)		
35–44	0.013 (0.135)	−0.445 (0.152)	0.432 (0.154)		
45–54	0.229 (0.124)	−0.213 (0.145)	−0.015 (0.161)		
54–64	0.003 (0.142)	0.152 (0.155)	−0.155 (0.172)		
≥65	−0.094 (0.169)	0.404 (0.181)	−0.310 (0.208)		
Educational level					
Primary education	0.015 (0.241)	0.470 (0.227)	−0.486 (0.334)	39.182	.000
Lower secondary education (vmbo)	0.072 (0.126)	0.418 (0.127)	−0.489 (0.168)		
Higher secondary education (havo/vwo)	−0.259 (0.142)	0.074 (0.156)	0.185 (0.160)		
Intermediate vocational training (mbo)	0.033 (0.129)	0.124 (0.139)	−0.157 (0.161)		
Higher vocational training (hbo)	0.020 (0.128)	−0.490 (0.140)	0.471 (0.141)		
University (wo)	0.119 (0.173)	−0.595 (0.219)	0.476 (0.188)		
Living arrangement					
Alone	0.033 (0.163)	−0.426 (0.191)	0.393 (0.203)	16.721	.033
With a partner, no children	−0.095 (0.153)	−0.174 (0.169)	0.269 (0.189)		
With a partner and children	−0.078 (0.146)	0.312 (0.164)	−0.234 (0.188)		
Alone with children	−0.045 (0.338)	0.284 (0.312)	−0.239 (0.447)		
Other	0.186 (0.374)	0.004 (0.436)	−0.189 (0.469)		
Paid job					
No	0.061 (0.078)	0.067 (0.087)	−0.128 (0.093)	1.920	.380
Yes	−0.061 (0.078)	−0.067 (0.087)	0.128 (0.093)		

The effect of the age-group variable is not significant. However, an analysis with exact age instead of the age-group variable, confirmed that a linear effect of age was significant for this first latent class, whereas no significant linear effect on the second and third latent class was found.

The general picture that emerges regarding the relationship between covariates and gender roles is that expressing a clear position—whether it is egalitarian or more gender stereotyping—is defined by the level of 'stability' in the lives of people. Respondents that experience uncertainty in their lives-i.e. single parents, not having a job, and being young-are more likely to be classified in the first latent class that show little differentiation in relative preference of the four gender role issue.

Example 3: Civil Morality

In the final example we apply our analyses to a set of six items referring to civil morality. Three of these items indicate justification of illegitimate personal enrichment, *that is*: 'claiming state benefits one is not entitled to'; 'cheating on tax'; and 'paying cash for services to avoid taxes'. The remaining items question the justification of interference in life and death of people, *i.e.* abortion, euthanasia and homosexuality. It is important to notice that these latter issues are legalized in Dutch society although political opinions on the matter differ. In most research (Halman, 1997) this distinction between the two sets of morality issues is maintained, supported by empirical evidence from factor analysis which identifies two dimensions. In this research, however, we aim at identifying latent classes that differ in their justification of these behaviours relative to an overall justification level. The latent-class regression model with random intercept suits this purpose.

The difference between models with and without random intercept is probably most clear-cut with this third example. Although the BIC criterion results in selecting a three class model in each case, the difference with a two class model in the traditional latent class regression model is small. Given that the estimated class size of the third latent class in this case is only 6 percent (compared to 18 percent in the model with random intercept), some researchers may conclude that a two class model isn't a bad choice. The differences in fit of the three class models with and without random intercept are also quite pronounced, both in terms of BIC and R^2. In the latter case the pseudo R^2 increases from .54 to .80 when adding a random intercept. This random intercept again strongly correlated with the variable which counted the number of agreement responses on the set of six items ($r = .855$).

The first and largest (54.5 percent) latent class appears to rank-order items in a 'Guttman fashion style' according to the difficulty of finding actions justifiable. 'Claiming state benefits' is the least and 'homosexuality' is the most likely behaviour to be justified. The three items regarding illegitimate personal enrichment rank lowest; the items on interference in life and death rank highest. However, the difference between 'paying cash' and 'abortion' is small. Rank-ordering within these two sets also seems to be logic when one reflects on current Dutch (and probably other Western) societies. Legislation on homosexuality is very liberal in the Netherlands, which was one of the first countries to legalize marriage among homosexual couples. Legislation on euthanasia and abortion are more strict but not prohibited. Given that euthanasia is a decision regarding one's own live and abortion involves the choice of a woman (or couple) on a premature conception, it seems obvious that public morality is more liberal in the case of euthanasia. None of the 'personal enrichment' issues is legal. However, there might be some confusion regarding the issue of 'paying cash for services to avoid taxes' since it is legal to pay somebody who can help you with optimizing your tax reductions. The major distinction, however, is between 'claiming state benefits' and the other two options. The former is regarded as "stealing from the poor" since most state benefits are provisions for the disadvantaged. When issues refer to "taxes" people associate it with the state or government and people tend to be less moralistic in their solidarity with the state.

The third latent class adopts a similar rank-ordering as the first latent class. The differences in beta's between items, however, are more pronounced - especially for the first and last item. Furthermore, the difference between 'paying cash for services' and 'abortion' is much more clear-cut than is the case for the first latent class. The second latent class, on the other hand, does not adopt a clear rank-ordering of the issues. In fact this category hardly differs from the overall justification except for two of the six items, *i.e.*, 'claiming state benefits' and 'paying cash for services'. Oddly, this category shows greater than average tolerance toward 'paying cash for services', while at the same time tend to find 'claiming state benefits' less justifiable. However, in comparison with the other two latent classes, this second class is the least likely to regard 'claiming state benefits' as not justifiable. This is also the case for the other two issues referring to personal enrichment suggesting that the second latent class has somewhat less moral objection against these illegitimate ways of self enrichment. Nevertheless, the dominant picture is that the latent profile of the second class does not strongly differ from the overall agreement pattern identified by the random intercept. As such it brings together a group of respondents who are virtually acting as non differentiators as far as their evaluation of moral behaviors is concerned.

Although the rank ordering in relative preference of the six items was highly similar for the first and third latent class, the effects of covariates in predicting class membership are distinct. Except for gender, none of the other covariates is significantly related to membership of the first latent class, as is indicated by the fact that beta's are less than twice their standard error.

Women are somewhat more likely than men to be classified in the first latent class. We already indicated that the rank ordering of the morality issues within the first latent class comes closest to what could be regarded as Dutch culture on these issues. Since the concept of a national culture implies a sharing of values or attitudes across different social groups within that society, it seems obvious that little effect of covariates could have been expected. Hence, the effect of covariates—or rather the lack of effect of covariates—on the likelihood of being classified in the first and largest latent class is merely supportive evidence for interpreting the latent profile of the first class as reflecting Dutch culture.

Covariates, however, are clearly related to classification in the third latent class. Women, higher educated respondents, respondents without children and respondents in the age group of 35–44 years is the emerging profile of this third latent class, as is indicated by their positive betas observed in Table 4. Recall that the rank ordering of items within this latent class is similar to the one observed in the first latent class, except that the differences are more pronounced as far as the first 'claiming state benefits' and last 'homosexuality' issues are concerned.

Are men more indifferent towards moral issues than women? That is the first question that pops in mind when we observe that men are more likely to be classified in the second latent class that hardly makes any difference in the rank ordering of items except for a moderately higher than overall agreement with the issue of 'paying cash for services to avoid taxes' and somewhat less agreement with 'claiming state benefits'. Also older respondents, aged 65 or more, and respondents with only primary or lower secondary education are more likely to be classified in this second latent class. In fact, the pattern of association between covariates and second class membership that did not substantially differentiate in their ranking of moral issues, appears to be the opposite pattern compared to the third latent class that did adopt a more pronounced item preference ranking. This opposite pattern of association suggests that the difference between the second and third latent class is a difference in strength by which a difference is made in justification of particular moral behaviours.

What Difference Does It Make?

In each of the three examples given we found that the latent class regression model with random intercept fitted the data better than the model without random intercept. Hence, in terms of measurement fit it does make a difference whether or not a random intercept is included to control for overall agreement tendencies in a set of items. An applied researcher, however, might wonder whether more substantive findings would change. He or she

TABLE 5 **Crosstabulation of predicted latent class membership (modal assignment) in the standard with the random intercept LC regression model**

Example 1: locus of control

		Standard LC regression		*Total*	
		LC1	*LC2*		
LC regression	LC1	766	0	766	72.9%
With random intercept	LC2	88	197	285	27.1%
	Total	854	197	1051	
		81.3%	18.7%		

Example 2: Gender role attitudes

		Standard LC regression			*Total*	
		LC1	*LC2*	*LC3*		
LC regression	LC1	595	0	0	595	56.6%
With random intercept	LC2	13	358	0	371	35.3%
	LC3	3	0	82	85	8.1%
	Total	611	358	82	1051	
		58.1%	34.1%	7.8%		

Example 3: Civil morality

		Standard LC regression			*Total*	
		LC1	*LC2*	*LC3*		
LC regression	LC1	372	218	3	593	56.4%
With random intercept	LC2	0	267	0	267	25.4%
	LC3	127	0	64	191	18.2%
	Total	499	485	67	1051	
		47.5%	46.1%	6.4%		

might wonder whether latent classes differ in content and/or size, or whether effects of covariates change. A lengthy discussion of similarities and differences is beyond the purpose of this article. We can put the answer very boldly: ‘yes it can and yes it does make a difference’. Let us illustrate this claim with some summary comparisons for the three examples given.

In none of the three examples there was clear evidence that the number of classes changes when including the random intercept. Class sizes and class membership, however, varied with distinct results in the three examples. The most parsimonious way of illustrating this is by cross-tabulating predicted class membership of the standard LC model with the random intercept model (Table 5). Class membership is defined by modal assignment given the latent class probabilities that are estimated by the model.

Class sizes differed when comparing the two types of models. Differences were smallest in the second 'gender role' example and most clearly observed in the third 'civil morality' example. The effect of adding a random intercept in the first 'locus of control' and second 'gender roles' example is that a number of people who were classified in the first latent class of the standard LC regression model switched to the second and/or third latent class in the random intercept model. In both cases this means that these respondents moved from a latent class in which latent class weights (betas) were moderate (or even insignificant) to a latent class in which the effect sizes were more pronounced. It is important to note that in both examples latent class weights were similar in the models with and without random intercept and that latent class probabilities of the corresponding latent classes correlated highly (>.90). Hence, in the first two examples the principal effect of introducing the random intercept was in fine-tuning the measurement model.

In the 'gender role' example, in which the random intercept could be interpreted as capturing agreement bias or acquiescence, the effects of covariates on latent class membership were very similar in the models with and without random intercept. The only differences we observed were that standard errors were generally somewhat less in the random intercept model and that effects of covariates were somewhat smaller. In the 'locus of control' example, however, the effect of education—which was the only significant covariate—was substantially smaller in the random intercept LC model compared to the standard LC regression model. In the latter model the difference in beta between the first and last category of education was equal to 2.045, whereas in the random intercept model this difference equaled 1.024 (see Table 2).

Finally, the third example on 'civil morality' revealed very different results when the standard LC model is compared with the random intercept model. Estimated latent class sizes differ substantially and one third of the respondents are classified in the off-diagonal cells of the table. Furthermore, latent class weights (betas) obtained with the traditional LC regression model also deviated from the betas reported in Table 4. Only the betas of the first latent class were similar. However, the latent class probability scores for this first class only correlated with .54, which is consistent with the previous finding that many respondents switched latent classes. Hence, it is safe to conclude that the measurement model of 'civil morality' is quite different when a random intercept is included. Estimates of the covariates were also very different, but given that we are talking about two truly different measurement models, a comparison of covariate effects makes little sense.

DISCUSSION

Our research demonstrates the usefulness of a method developed in the context of consumer research, that is, a latent-class ordinal regression model with random intercept, in controlling for overall agreement with rating questions in public opinion research. We argued that the desire to control for such overall agreement may arise from a substantive motive, *i.e.*, when a researcher's conceptual framework refers to relative preferences of particular issues compared to others rather than absolute agreement rating. Controlling for overall agreement can also be justified from a methodological point of view when a balanced (positively and negatively worded) set of items is analyzed. In this case overall agreement can be interpreted as agreement bias. The method is applicable when a researcher is interested in developing an empirical typology. As such it does not 'replace' other methods developed to control for overall agreement response patterns when a researcher's aim is to identify latent dimensions. To our knowledge, research on overall agreement response patterns in public opinion research has been exclusively developed within this dimensional—factor analytic type of—perspective. The method used in this research is a valuable addition for researchers working within the typology—cluster analytic type of—perspective. Furthermore, the approach also allows for inclusion of covariates predicting latent class membership. As such the model fits within the structural equation modeling framework. With software readably available and given the examples from regular public opinion surveys that were presented, we hope that public opinion researchers will consider the use of this approach in their own work when they share the concern that guided this research, *i.e.*, controlling for overall agreement response tendencies in the context of identifying segments or clusters in a population that differ in their relative preferences of particular items compared to other items in the response set.

REFERENCES

Agresti, A. (2002). *Categorical data analysis* (2nd edn) New York: Wiley.

Alwin, D.F., & Krosnick, J. A. (1988). A test of the form resistant correlation hypothesis – Ratings, rankings and the measurement of values. *Public Opinion Quarterly, 52*, 526–538.

Billiet, J., & McClendon, M. J. (2000). Modeling acquiescence in measurement models for two balanced sets of items. *Structural Equation Modeling, 7*, 608–628.

Cattell, R. B. (1944). Psychological measurement: normative, ipsative, interactive. *Psychological Review, 44*, 292–303.

Cheung, M. W.-L., & Chan, W. (2002). Reducing uniform response bias with ipsative measurement in multiple-group confirmatory factor analysis. *Structural Equation Modeling, 9*, 55–77.

Cheung, M. W.-L. (2006). Recovering preipsative information from additive ipsative data: a factor score approach. *Educational and Psychological Measurement, 66*, 565–588.

Cunningham, W. H., Cunningham, I. C. M., & Green, R. T. (1977). The ipsative process to reduce response set bias. *The Public Opinion Quarterly, 41*, 379–384.

Dunlap, W. P., & Cornwell, J. M. (1994). Factor analysis of ipsative measures. *Multivariate Behavioral Science, 29*, 115–126.

Eid, M., Langeheine, R., & Diener, E. (2003). Comparing typological structures across cultures by multigroup latent class analysis. A primer. *Journal of Cross-Cultural Psychology, 34*, 195–210.

Halman, L. C. J. M. (1997). Is there a moral decline? A cross national inquiry into morality in contemporary society. *International Social Science Journal, 13*, 49–69.

Inglehart, R. (1990). *Culture shift in advanced industrial society*. Princeton University Press, Princeton, NJ.

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in data: An introduction to cluster analysis*. New York: John Wiley & Sons, Inc.

Klein, M., Dülmer, H., Ohr, D., Quandt, M., & Rosar, U. (2004). Response sets in the measurement of values: A comparison of rating and ranking procedures. *International Journal of Public Opinion Research, 16*, 474–483.

Kohn, M., & Schooler, C. (in cooperation with Miller, J., Miller, K., Schoenbach, C. & Schoenberg, R.) (1983). *Work and personality. An inquiry into the impact of social stratification*. Norwood, New Jersey: Ablex Publishing Corporation.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 140*, 1–55.

Magidson, J., & Vermunt, J. K. (2006). Use of latent class regression models with a random intercept to remove overall response level effects in ratings data. In: A. Rizzi & M. Vichi (Eds.), *Proceedings in Computational Statistics* (pp. 351–360). Heidelberg: Springer.

Mirowsky, J., & Ross, C. E. (1990). Control or defense? Depression and the sense of control over good and bad outcomes. *Journal of Health and Social Behavior, 31*, 71–86.

Munson, J. M., & McIntyre, S. H. (1979). Developing practical procedures for the measurement of personal values in cross-cultural marketing. *Journal of Marketing Research, 16*, 48–52.

Paulhus, D. L. (1991). Measures of personality and social psychological attitudes. In: J. P. Robinson & R. P. Shaver (Eds.), *Measures of Social Psychological Attitudes Series* (Vol. 1, pp. 17–59). San Diego: Academic.

Popper, R., Kroll, J., & Magidson, J. (2004). Application of latent class models to food product development: a case study. *Sawtooth Software Conference Proceedings*. San Diego, CA: Sawtooth Software, Inc.: Sequim, WA.

Ray, J. J. (1979). Is the acquiescent response style problem not so mythical after all?: Some results from a successful balanced F scale. *Journal of Personality Assessment, 43*, 638–643.

Rotter, J. B. (1954). *Social learning and clinical psychology*. New York: Prentice-Hall.

Ryan, M., Bate, A., Eastmond, C. J., & Ludbrook, A. (2001). Use of discrete choice experiments to elicit preferences. *Quality in Health Care, 10*(Suppl 1), i55–i60.

Thornton, A., & Young-DeMarco, L. (2001). Four decades of trends in attitudes toward family issues in the United States: The 1960s through the 1990s. *Journal of Marriage and the Family, 63*, 1009–1037.

van Herk, H., & van de Velden, M. (2007). Insight into the relative merits of rating and ranking in a cross-national context using three-way correspondence analysis. *Food Quality and Preference, 18*, 1096–1105.

Vermunt, J. K., & Magidson, J. (2005). *Latent GOLD 4.0 User's Guide*. Belmont Massachussetts: Statistical Innovations Inc.

BIOGRAPHICAL NOTE

Guy Moors is Assistant Professor at the Department of Methodology & Statistics, Faculty of the Social and Behavioural Sciences, Tilburg University. His research interests are in the field of social demography, values and attitude research, survey research methodology and latent class analysis.