

# How to perform three-step latent class analysis in the presence of measurement non-invariance or differential item functioning

Jeroen K. Vermunt<sup>1</sup> and Jay Magidson<sup>2</sup>

<sup>1</sup>Tilburg University; <sup>2</sup>Statistical Innovations Inc.

## ABSTRACT

The practice of latent class (LC) modeling using a bias-adjusted three-step approach has become widely popular. However, the current three-step approach has one important drawback – its key assumption of conditional independence between external variables and latent class indicators is often violated in practice, such as when a (nominal) covariate represents subgroups showing measurement non-invariance (MNI) or differential item functioning (DIF). In this article, we demonstrate how the current three-step approach should be modified to account for MNI; that is, covariates causing DIF should be included in the step-one model and the step-three classification error adjustment should differ across the values of the DIF covariates. We also propose a model-building strategy that makes the new methodology practically applicable also when it is unknown which of the external variables cause DIF. The new approach, implemented in the program Latent GOLD, is illustrated using a synthetic and a real data example.

## KEYWORDS

auxiliary variables; item bias; mixture modeling; three-step modeling

## Introduction

During the past several years, the practice of latent class (LC) modeling using a stepwise approach called bias-adjusted three-step LC analysis has become widely popular. It originates from the work of Bolck et al. (2004), who proposed a stepwise approach for modeling the association between classes and external variables. This involves:

- (1) Estimating the LC measurement model parameters, including the determination of the number of latent classes.
- (2) Classifying cases into one of the classes based on the model selected in step one (typically, the most likely class).
- (3) Examining the relationship between the classes and external or auxiliary variables while accounting for classification errors introduced in step two.

The external variables used in step three can be covariates affecting the classes (Vermunt, 2010), distal outcomes affected by the classes (Bakk & Vermunt, 2016), or a combination of these (Bakk et al., 2013; Nylund-Gibson et al., 2019). Also, more complex models can be estimated in the third step, such as latent transition or latent Markov models (Asparouhov & Muthén, 2014; Di Mari et al., 2016). The trend toward this stepwise approach to deal with covariates and distal outcomes is due to its practical advantages, its several theoretical benefits, and its implementation in generally available LC analysis software like Mplus (Muthén & Muthén, 2015) and Latent GOLD (Vermunt & Magidson, 2016).

An alternative approach, referred to as one-step LC modeling, involves simultaneous inclusion of class indicators and

external variables in a single model (Bandein-Roche et al., 1997; Dayton & Macready, 1988). For a list of problems in this one-step approach that are resolved by the three-step approach, as well as a description of the main three-step approaches, see Vermunt (2010). See also Nylund-Gibson and Masyn (2016), who present a simulation study that reveals several problems with respect to the selection of the number of classes when covariates are included early on in the modeling process.

However, despite its many improvements over the one-step approach, the current three-step LC analysis approach has one important drawback – its key assumption of conditional independence between external variables and latent class indicators is often violated in practice (Masyn, 2017). More specifically, the derivation of the bias-adjusted three-step method requires the assumption of no direct effects between external variables (covariates) and the response variables used to construct the LC model (indicators). Violations of this assumption include situations where a covariate represents subgroups showing measurement non-invariance (MNI) or item bias (Clogg & Goodman, 1984; Kankaras, Moors, Vermunt, 2010). But, more generally, the LC model with covariates is a multiple indicators multiple causes (MIMIC) model where, as in factor analysis and item response theory models, DIF can be conceptualized as causes having not only indirect effects on the indicators via the latent variable(s) but also having direct effects on the indicators (Lee et al., 2017; Suh & Cho, 2014; Woods, 2009). In the remainder, we use the terms MNI, DIF, and covariates/causes with direct effects interchangeably.

A simulation study designed by Asparouhov and Muthén (2014) to investigate the consequences of ignoring direct effects of covariates on indicators demonstrated that the three-step

approach may yield seriously biased estimates in such a case (see also Janssen et al., 2019). As a possible solution, they presented a somewhat ad-hoc modification of the three-step approach in which the direct effects of interest are included in the measurement part of the step-one model. A similar approach was used by Di Mari and Bakk (2018) as a modification of two-step and three-step latent Markov modeling with direct effects. While these approaches perform better than fully ignoring the direct effects, a certain amount of bias remains, which may still be substantial.

Based on the above, one may conclude three-step LC analysis should not be performed in situations with direct effects, and a one-step model should be used instead. In fact, Janssen et al. (2019) did not even attempt to include the three-step approach in their simulation study in the detection and modeling of direct effects in LC analysis. They state: “With three-step, in contrast, adding DEs [Direct Effects] becomes increasingly difficult and often even impossible.” However, as we demonstrate in this article, there turns out to be a rather simple way to modify the three-step approach to allow for direct effects. Specifically, the correct three-step approach with MNI or DIF proceeds as follows:

- (1) Covariates with direct effects on the indicators are included in the step-one model, whereas other external variables need not be included. The step-one model should contain not only the relevant direct effects but also the effects of the included covariates on the latent classes.
- (2) The step-two classifications are obtained in the usual way, but now also account for the fact that these depend on the values of the covariates included in the step-one model.
- (3) The step-three model of interest contains the DIF covariates used in step one, as well as other covariates and/or distal outcomes of interest. The key modification here compared to the standard step-three modeling is that the classification error correction matrix, also referred to as D matrix, should be allowed to differ across categories of covariates with direct effects.

One natural application of the modified step-three approach is in situations where it is desired to account for MNI, such as in cross-cultural research. In the step-one analysis, one includes the country as a grouping variable and builds a measurement model that may differ in certain ways across countries (that is, one or more indicators may be affected by DIF across countries). The classification in step two is based on all variables used in the step-one model, including the country. The step-three model would include country as one of the covariates and account for the fact that classification errors (the D matrices) differ across countries. This method of dealing with direct effects has been implemented in Latent GOLD 6.0, in both the Step3 and the Syntax module (Vermunt & Magidson, 2020).

Note that our modified approach differs in two important ways from the modification Asparouhov and Muthén (2014) experimented with and Di Mari and Bakk (2018) used in the context of latent Markov modeling. First, the step-one model

contains the effects of the covariates on the classes, which prevents the overestimation of the effects due to the omission of the indirect effects via the classes. Second, the step-three correction for classification errors differs across categories of the external variables with direct effects, which is exactly what is required to properly account for the non-invariant measurement models and avoid biasing the model parameters.

The next section provides the theoretical framework for our modified three-step procedure and explains how to apply the new approach in practice, that is, when we do not know a priori which external variables are causing DIF. Following that, we illustrate the new approach with both synthetic and real data applications. In Appendix A1, we show how to set up the models using the Latent GOLD 6.0 software. We end with a conclusion and discussion section.

### Three-step LC analysis with direct effects

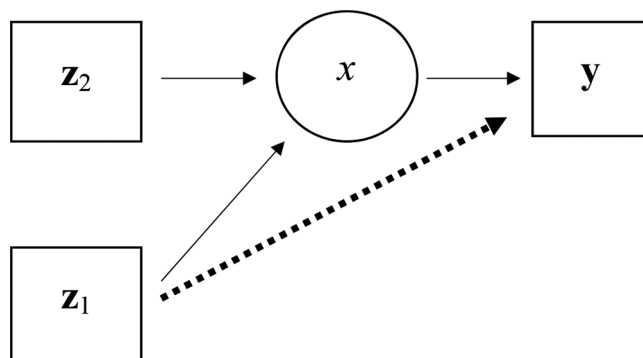
In this section, we derive the form of the step-one and step-three models for three-step LC analysis with direct effects. We consider three different situations where direct effects need to be addressed in somewhat different ways: 1) covariates have direct effects on indicators, 2) covariates have direct effects on both indicators and distal outcomes, and 3) indicators have direct effects on distal outcomes. We also propose a model-building strategy that can be used when it is not known a priori which covariates may induce DIF.

#### Covariates with direct effects on indicators

The population model of interest is an LC model with two sets of explanatory variables,  $\mathbf{z}_1$  and  $\mathbf{z}_2$ , where the  $\mathbf{z}_1$  have direct effects on some of the response variables or indicators  $\mathbf{y}$ , whereas the  $\mathbf{z}_2$  do not. This situation is depicted in Figure 1. The LC model for  $P(\mathbf{y}|\mathbf{z}_1, \mathbf{z}_2)$ , the probability of having a certain pattern of responses conditional on the covariate values, takes on the following form:

$$P(\mathbf{y}|\mathbf{z}_1, \mathbf{z}_2) = \sum_x P(x|\mathbf{z}_1, \mathbf{z}_2)P(\mathbf{y}|x, \mathbf{z}_1) \quad (1)$$

Here,  $P(x|\mathbf{z}_1, \mathbf{z}_2)$  is the probability of belonging to latent class  $x$  conditional on the covariate values, which will usually be modeled using a logistic regression equation. The term



**Figure 1.** LC model with covariates  $\mathbf{z}_1$  and  $\mathbf{z}_2$  and class indicators  $\mathbf{y}$ . The modified three-step approach is needed when covariates  $\mathbf{z}_1$  have direct effects on one or more of the indicators  $\mathbf{y}$  (see thick dashed line).

$P(\mathbf{y}|x, \mathbf{z}_1)$  denotes the probability of a response pattern  $\mathbf{y}$  given class membership  $x$  and covariate values  $\mathbf{z}_1$ . Under the typical assumption of local independence between the response variables, we obtain:

$$P(\mathbf{y}|x, \mathbf{z}_1) = \prod_j P(y_j|x, \mathbf{z}_1). \quad (2)$$

Here,  $P(y_j|x, \mathbf{z}_1)$  is the class-specific response probability for the  $j$ th indicator, which, because of the direct effects, may also depend on  $\mathbf{z}_1$ . For categorical responses, these probabilities will usually be modeled using logistic equations.

Following the same logic as a standard three-step LC analysis, since  $\mathbf{z}_2$  does not affect  $\mathbf{y}$  directly, we can ignore  $\mathbf{z}_2$  in the estimation of the measurement part of the model. The resulting step-one model for  $P(\mathbf{y}|\mathbf{z}_1)$  is obtained as follows:

$$\begin{aligned} P(\mathbf{y}|\mathbf{z}_1) &= \sum_{\mathbf{z}_2} P(\mathbf{z}_2|\mathbf{z}_1)P(\mathbf{y}|\mathbf{z}_1, \mathbf{z}_2) \\ &= \sum_x \sum_{\mathbf{z}_2} [P(\mathbf{z}_2|\mathbf{z}_1)P(x|\mathbf{z}_1, \mathbf{z}_2)]P(\mathbf{y}|x, \mathbf{z}_1) \\ &= \sum_x P(x|\mathbf{z}_1)P(\mathbf{y}|x, \mathbf{z}_1), \end{aligned} \quad (3)$$

that is, by collapsing or marginalizing over the  $\mathbf{z}_2$ . As can be seen, the term of main interest,  $P(\mathbf{y}|x, \mathbf{z}_1)$ , is still the same as in the model that includes  $\mathbf{z}_2$ , which shows that we can ignore  $\mathbf{z}_2$  in the step-one model. It also shows that 1) the covariates  $\mathbf{z}_1$  should be part of the step-one model, 2) the specification used for  $P(\mathbf{y}|x, \mathbf{z}_1)$  should be the same as in the full model, and 3) the  $\mathbf{z}_1 - x$  associations should be included in the step-one model. Note that the marginal  $\mathbf{z}_1 - x$  associations in  $P(x|\mathbf{z}_1)$  will usually not be the same as the partial  $\mathbf{z}_1 - x$  association in  $P(x|\mathbf{z}_1, \mathbf{z}_2)$ .

Using the above step-one model, in step two, one can obtain class assignments  $w$  based on the posterior class membership probabilities,  $P(x|\mathbf{y}, \mathbf{z}_1) = P(x|\mathbf{z}_1)P(\mathbf{y}|x, \mathbf{z}_1)/P(\mathbf{y}|\mathbf{z}_1)$ . When using modal assignment,  $P(w|\mathbf{y}, \mathbf{z}_1)$  equals 1 for the class with the largest  $P(x|\mathbf{y}, \mathbf{z}_1)$  and 0 for the other classes. With proportional assignment,  $P(w|\mathbf{y}, \mathbf{z}_1)$  and  $P(x|\mathbf{y}, \mathbf{z}_1)$  are equal to one another. Irrespective of the class assignment method that is used, the relationship between the assigned class  $w$  and the true class  $x$ , which is crucial for the step-three adjustment, is obtained as follows:

$$P(w|x, \mathbf{z}_1) = \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{z}_1)P(x|\mathbf{y}, \mathbf{z}_1)P(w|\mathbf{y}, \mathbf{z}_1)/P(x|\mathbf{z}_1), \quad (4)$$

where for  $P(\mathbf{y}|\mathbf{z}_1)$  we can use either the estimated or the empirical distribution. Note that this formula is the same as the one provided by Bolck et al. (2004) and Vermunt (2010), except for the conditioning on  $\mathbf{z}_1$ .

As shown by Bolck et al. (2004), to derive the step-three model for  $P(w|\mathbf{z}_1, \mathbf{z}_2)$ , we first define the structure for the joint distribution  $P(x, \mathbf{y}, w|\mathbf{z}_1, \mathbf{z}_2)$  based on the assumed population model and the classification rule, that is,

$$P(x, \mathbf{y}, w|\mathbf{z}_1, \mathbf{z}_2) = P(x|\mathbf{z}_1, \mathbf{z}_2)P(\mathbf{y}|x, \mathbf{z}_1)P(w|\mathbf{y}, \mathbf{z}_1) \quad (5)$$

which is subsequently collapsed over  $x$  and  $\mathbf{y}$  to obtain  $P(w|\mathbf{z}_1, \mathbf{z}_2)$ :

$$\begin{aligned} P(w|\mathbf{z}_1, \mathbf{z}_2) &= \sum_x P(x|\mathbf{z}_1, \mathbf{z}_2) \sum_{\mathbf{y}} P(\mathbf{y}|x, \mathbf{z}_1)P(w|\mathbf{y}, \mathbf{z}_1) \\ &= \sum_x P(x|\mathbf{z}_1, \mathbf{z}_2)P(w|x, \mathbf{z}_1). \end{aligned} \quad (6)$$

The resulting step-three model is a standard LC model with a single indicator  $w$  with “response” probabilities  $P(w|x, \mathbf{z}_1)$ , which were defined above. Bolck et al. (2004) called the matrix with elements  $P(w|x)$  the D matrix. This matrix makes it possible to correct for the classification errors in the assigned class memberships  $w$ . The key result of our derivation for the DIF case is that the D matrix should be allowed to vary across levels of the DIF covariates. Estimation of the step-three model yields the effects of both the  $\mathbf{z}_1$  and the  $\mathbf{z}_2$  variables on the true latent classes  $x$ , which appear in the term  $P(x|\mathbf{z}_1, \mathbf{z}_2)$ . For parameter estimation, one can use either the method proposed by Bolck et al. (2004), which is often referred as the BCH approach or the maximum likelihood approach suggested by Vermunt (2010).

### Covariates with direct effects on both indicators and distal outcomes

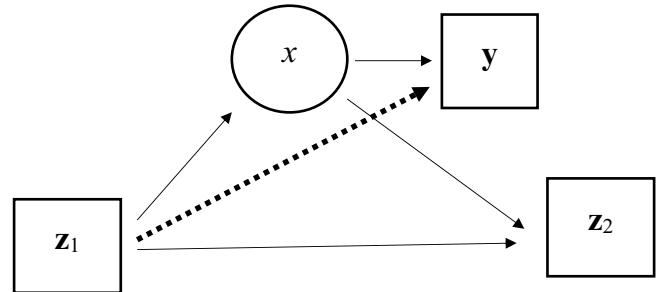
Suppose the population model of interest is a model in which the  $\mathbf{z}_2$  are distal outcomes instead of covariates. As illustrated in Figure 2, the covariates  $\mathbf{z}_1$  affect both the class indicators and the distal outcomes directly; that is, they may cause DIF and are at the same time relevant control variables in the model for the distal outcomes. In this case, the underlying population model becomes:

$$P(\mathbf{y}, \mathbf{z}_2|\mathbf{z}_1) = \sum_x P(x|\mathbf{z}_1)P(\mathbf{y}|x, \mathbf{z}_1)P(\mathbf{z}_2|x, \mathbf{z}_1). \quad (7)$$

Because the  $\mathbf{z}_2$  are unrelated to the  $\mathbf{y}$  conditional on  $x$  and  $\mathbf{z}_1$ , we can use the same step-one model and step-two classification as described above for the covariate case. The step-three model will have the following form:

$$P(w, \mathbf{z}_2|\mathbf{z}_1) = \sum_x P(x|\mathbf{z}_1)P(w|x, \mathbf{z}_1)P(\mathbf{z}_2|x, \mathbf{z}_1). \quad (8)$$

This shows we should include  $\mathbf{z}_1$  in the step-three model, and, as in the covariate case, use classification error probabilities  $P(w|x, \mathbf{z}_1)$  which depend on  $\mathbf{z}_1$ .



**Figure 2.** LC model with covariates  $\mathbf{z}_1$ , class indicators  $\mathbf{y}$ , and distal outcomes  $\mathbf{z}_2$ . Covariates  $\mathbf{z}_1$  are also control variables in the model for the distal outcomes  $\mathbf{z}_2$ . The modified three-step approach is needed when covariates  $\mathbf{z}_1$  have direct effects on one or more of the indicators  $\mathbf{y}$  (see thick dashed line).

In the simpler case in which  $\mathbf{z}_2$  is not affected by  $\mathbf{z}_1$  directly, the step-one model remains the same. The step3-model can be simplified to:

$$P(w, \mathbf{z}_2) = \sum_x P(x)P(w|x)P(\mathbf{z}_2|x), \quad (9)$$

where  $P(w|x)$  is computed in the usual manner, that is, by aggregating over both  $\mathbf{y}$  and  $\mathbf{z}_1$ .

### Indicators with direct effects on distal outcomes

Suppose the population model of interest is a model in which some of the responses  $\mathbf{y}$  (denoted by  $\mathbf{y}_1$ ) have a direct effect on distal outcomes  $\mathbf{z}$  (see Figure 3). In this case, the underlying population model becomes:

$$P(\mathbf{y}_1, \mathbf{y}_2, \mathbf{z}) = \sum_x P(x)P(\mathbf{y}_1|x)P(\mathbf{y}_2|x)P(\mathbf{z}|x, \mathbf{y}_1). \quad (10)$$

Because  $\mathbf{z}$  follows causally after  $\mathbf{y}_1$ , despite the direct effect, we can use a standard step-one model that excludes the distal outcomes. The step-three model will have the following form:

$$P(w, \mathbf{z}, \mathbf{y}_1) = \sum_x P(x)P(w, \mathbf{y}_1|x)P(\mathbf{z}|x, \mathbf{y}_1). \quad (11)$$

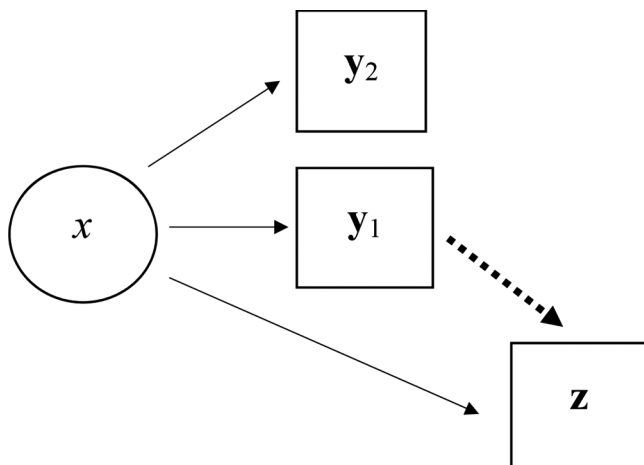
Thus, in the step-three model, we should again include the  $\mathbf{y}_1$  while accounting for its relationship with  $w$  and  $x$ . This model may also be written as a model conditional on  $\mathbf{y}_1$ :

$$P(w, \mathbf{z}|\mathbf{y}_1) = \sum_x P(x|\mathbf{y}_1)P(w|x, \mathbf{y}_1)P(\mathbf{z}|x, \mathbf{y}_1), \quad (12)$$

which implies using  $\mathbf{y}_1$  in a similar way as  $\mathbf{z}_1$  was used in the step-three models described above.

### A three-step model-building strategy

The three-step approach described above requires the inclusion of the covariates inducing MNI or DIF in the step-one model. However, in practice, we often do not know a priori which of the covariates of interest will have direct effects on the



**Figure 3.** LC model with class indicators  $\mathbf{y}_1$  and  $\mathbf{y}_2$  and distal outcomes  $\mathbf{z}$ . The modified three-step approach is needed when class indicators  $\mathbf{y}_1$  have direct effects on the distal outcomes  $\mathbf{z}$  (see thick dashed line).

indicators. To overcome this problem, we propose using the following model-building strategy:

- (1) Construct the step-one LC model:
  - a. Determine the number of classes and other relevant LC model features without the inclusion of covariates.
  - b. For each covariate separately, include it in the model selected in step 1a and determine whether it has direct effects on the indicators.
  - c. Specify the final step-one model with all DIF covariates and the significant direct effects encountered in step 1b.
- (2) Obtain the classifications.
- (3) Estimate the step-three model with all covariates and/or distal outcomes, taking into account that classification errors depend on the DIF covariates.

Step 1a is based on the work by Nylund-Gibson and Masyn, who showed that especially in the presence of DIF, class enumeration should be done using the class indicators only, thus without covariates (or distal outcomes). For step 1b, it is key to have a tool to identify the direct effects, such as Bivariate Residuals reported by the Latent GOLD program (Oberski et al., 2013; Vermunt & Magidson, 2016). When estimating the LC model with DIF from step 1c, the LC analysis software will allow the user to write the classifications to a file (step 2).

While steps 1a, 1b, 1c, and 2 can be performed with any LC analysis software, step 3 requires an option to indicate which are the DIF covariates. Appendix A1 shows how this is implemented in Latent GOLD version 6.0 (Vermunt & Magidson, 2020).

### Two example illustrations

#### A demonstration using a constructed data example

To show that the proposed three-step approach with DIF works, we will apply the method to a synthetic data set. That is, we will demonstrate the new method recovers the parameters from a known population perfectly, while the standard approach does not, which confirms that the standard approach yields biased estimates. The population model we used consists of three latent classes, six dichotomous indicators ( $\mathbf{y}_1$ - $\mathbf{y}_6$ ), and two covariates ( $\mathbf{z}_1$  and  $\mathbf{z}_2$  having two categories,  $-1$  and  $1$ , and  $\mathbf{z}_2$  with five categories,  $-2, -1, 0, 1,$  and  $2$ ). The specified LC population model yielded  $P(\mathbf{y}|\mathbf{z}_1, \mathbf{z}_2)$ , which we multiplied by the number of observations of the covariate pattern concerned. The total “sample” size was set to 1000. The resulting non-integer frequencies were used as frequency weights in the LC analysis of the construct data set.

The values used for the class-specific response probabilities were the same as those used by Vermunt (2010) and Bakk et al. (2013) for the moderate class separation condition; that is, the “success” probability for Class 1 equals .80 on all indicators, for Class 2 .20 on all indicators, and for Class 3 .80 on the first three indicators and .20 on the other three. These “success” probabilities correspond with logit values of 1.3863 and  $-1.3863$ , respectively. The direct effects of  $\mathbf{z}_1$  on  $\mathbf{y}_1$  and  $\mathbf{y}_4$  were set to 1, implying a shift of  $-1$  and  $1$  in the response logit for the first and second



categories. of  $z_1$ , respectively. The intercept,  $z_1$  slope, and  $z_2$  slope parameters in the logistic model for the classes were set to  $-1$ ,  $-1$ , and  $1$  for Class 2 and to  $-2$ ,  $1$ , and  $-1$  for Class 3 (Class 1 served as baseline category). The resulting class proportions were .53, .31, and .16. The entropy-based R-squared value was .71 with covariates and .60 without covariates, corresponding with a moderate class separation.

Table 1 reports the parameter values of the regression model for the latent classes obtained using different LC analysis approaches, where for the three-step methods modal class assignment and maximum likelihood bias-adjustment was used. As can be seen, both the correctly specified one-step model and the new three-step approach accounting for MNI or DIF reproduce the true parameter values perfectly. In contrast, the one-step model without DIF, the three-step approach without DIF, and the three-step approach with DIF in step one but not in step-three yield severely biased estimates for the  $z_1$  slopes. The intercept and  $z_2$  slope estimates are close to their true values.

In this synthetic example, the absolute bias in the  $z_1$  effects turned out to be about 30% for each of the “incorrect” approaches. However, it should be noted that in practice the amount of bias will depend on factors such as strength and direction of the DIF, number of DIF items, and DIF being uniform or non-uniform. (Janssen et al., 2019; Masyn, 2017). We noticed that even the settings used in the step-three model estimation matter: the amount of bias changed with proportional instead of a modal class assignment or with BCH instead of maximum likelihood adjustment.

### An illustration with a real data example

The extended three-step approach for dealing with MNI or DIF and the proposed model-building strategy will be illustrated with a data set used by Hagenaars (1993) in his book on

**Table 1.** Population Values and Parameter Estimates Obtained Using Various LC Approaches.

Model	Class 2			Class 3		
	Intercept	$z_1$	$z_2$	Intercept	$z_1$	$z_2$
Population model	-1.000	-1.000	1.000	-2.000	1.000	-1.000
One-step LC analysis with DIF	-1.000	-1.000	1.000	-2.000	1.000	-1.000
Three-step LC analysis with DIF	-1.000	-1.000	1.000	-2.000	1.000	-1.000
One-step LC analysis without DIF	-1.025	-1.318	1.014	-2.017	0.706	-1.020
Three-step LC analysis without DIF	-1.006	-1.276	1.014	-2.023	0.718	-1.017
Three-step LC analysis with DIF in step-one only	-1.004	-1.270	1.010	-2.025	0.718	-1.018

The three-step analyses are based on modal classification and maximum likelihood bias adjustment.

**Table 3.** Bivariate Residuals for the Covariates Obtained by Including Them One at a Time in the Selected Step 1a Model, the Confirmatory 2-DFactor Model (This is Analysis Step 1b).

Covariate	Indicator				
	System responsiveness	Ideological level	Repression potential	Protest approval	Conventional participation
Gender	2.818	2.578	0.001	0.027	0.026
Education	0.202	2.307	0.005	0.000	0.590
Age	4.350	0.555	4.530	1.316	0.953

**Table 2.** Fit Statistics for the LC Models Estimated with the Data Set from Hagenaars (1993) Using Only the Class Indicators (This is Analysis Step 1a).

Model	Log-likelihood	BIC	AIC	L2	df	p-value
Unrestricted 1-class	-3767.25	7569.77	7544.51	296.56	26	<0.001
Unrestricted 2-class	-3666.88	7411.35	7355.77	95.82	20	<0.001
Unrestricted 3-class	-3631.38	7382.65	7296.75	24.80	14	0.036
Unrestricted 4-class	-3622.70	7407.61	7291.40	7.45	8	0.49
Exploratory 2-DFactor	-3625.13	7370.15	7284.25	12.30	14	0.58
Confirmatory 2-DFactor	-3628.70	7349.09	7283.40	19.45	18	0.36

loglinear modeling with latent variables. It is the demo data file “political.sav” which is part of the Latent GOLD software package. It contains five dichotomous indicators measuring what Hagenaars referred to as “System involvement” and “Protest tolerance,” as well as the categorical covariates’ gender, education (college and less than college), and age (18–34, 35–57, and 58–91).

Table 2 reports the fit measures for the estimated LC models in step 1a (without covariates). These indicate that according to the BIC we should select a 3-class model and according to the AIC and the likelihood-ratio goodness-of-fit statistic a 4-class model. The other two models are restricted 4-class models in which the classes are “forced” to represent two dichotomous latent variables. Such models are sometimes referred to as a discrete factor (DFactor) model (Magidson & Vermunt, 2001). These two restricted models are preferred over the unrestricted 3- and 4-class models, where the “confirmatory” variant gives the best balance between fit and parsimony. In the latter model, three indicators load on the “System involvement” factor and two on the “Protest tolerance” factor. This model was also selected as the best model by Hagenaars (1993).

In step 1b, we estimated three models which included one of the covariates in the model selected in step 1a. Table 3 presents the Bivariate Residuals reported by Latent GOLD, which are approximate chi-squared statistics with 1 degree of freedom. Using a cutoff value of 3, we conclude there is evidence for age having direct effects on two of the indicators, while for gender and education we do not need direct effects. In step 1c, we, therefore, use a model with age as a covariate and the two encountered direct effect included. This model is also used for classification (step 2).

Table 4 presents the parameter estimates for the step-three model which accounts for DIF, and, for comparison, also for the standard three-step approach which ignores DIF (these are both based on proportional classification and maximum likelihood bias adjustment, the default setting in Latent GOLD). Rather than estimating a single model for the four classes (the four combinations of the two dichotomous latent variables), we performed separate step-three analyses (binary logit regression analyses) for “System involvement” and “Protest tolerance”,

**Table 4.** Parameter Estimates of the Logistic Regression Models for System Involvement and Protest Approval Obtained Using Three-Step Analyses with and without DIF.

	Model for system involvement				Model for protest tolerance			
	DIF		No-DIF		DIF		No-DIF	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Intercept	1.560	0.228	2.394	0.285	1.997	0.194	2.086	0.201
Female	-0.710	0.209	-0.702	0.223	-0.261	0.159	-0.282	0.161
Less than college	-2.758	0.270	-2.888	0.293	-1.224	0.163	-1.232	0.165
Age 35–57	1.516	0.302	0.758	0.280	-0.904	0.189	-0.975	0.192
Age 58–91	0.806	0.293	0.047	0.280	-1.924	0.200	-2.015	0.206

Male, College, Age 18–34, Low System Involvement, and Low Protest Approval are the reference categories. Proportional class assignment and maximum likelihood bias-adjustment was used.

with “low” used as the reference class. As can be seen, accounting for the DIF induced by age has a huge impact on the estimated age effects on “System involvement” but only a small impact on its effects on “Protest tolerance.” The latter can be explained from the fact that the DIF item loading on the second factor (Repression potential) is not very strongly related to this factor, implying it does not have a strong impact on the classification and its errors. Content wise, we can conclude that Male, College, and Age 35–57 or 58–91 are associated with higher “System involvement” and College and Age 18–34 are associated with higher “Protest approval.”

## Conclusion and discussion

In this paper, we derived the correct modification of the current three-step LC analysis approaches that need to be made to properly adjust for the situation in which one of its basic assumptions is violated. The three situations discussed concerned: covariates with direct effects on indicators, covariates with direct effects on both indicators and distal outcomes, and indicators having direct effects on distal outcomes. For the first situation, we used a synthetic data example to demonstrate the new method works by applying it to a hypothetical population where DIF is present. In addition, a model-building strategy was presented which can be used to find the external variables causing DIF, thus making the new methodology practically applicable. The approach was illustrated with a real-life application using a data set from Hageñaars (1993), which showed that using the new approach may have a large impact on the step-three parameter estimates.

It is important to note that correct modeling of MNI or DIF is not only necessary for a valid step-three LC analysis but it can also be of interest on its own merits. For example, in cross-national comparative research, the step-one LC model will typically include country as a covariate or grouping variable, where direct effects are included to account for country differences in the interpretation of the items. Individuals can then be classified based on the resulting multiple-group LC model which accounts for country differences (Clogg & Goodman, 1984; Kankaras, Moors, Vermunt, 2010; McCutcheon, 1987). However, this does not mean that country needs to be included in a step-three LC analysis. Alternatively, a step-three analysis may involve the prediction of class membership using individual-level predictors only (say age, education, and gender). In such a case, a standard step-three analysis should be performed in which one does not account for MNI across countries. That is, only if the country is part of the step-three model, country DIF should be accounted for in this model.

It is important to note our paper assumes that indicators affected by DIF are to be retained in the step-one LC analysis. An alternative approach to handle the presence of DIF is to remove the DIF items from the measurement model, which is commonly done in the field of item response theory modeling. However, because LC analyses are usually performed with a relatively small number of indicators, one will typically prefer keeping *all* indicators, which requires modeling the MNI in the right way in the step-one model, and subsequently using the new approach described in this paper in the step-three model.

## Limitations

The proposed method works well when the number of covariates or grouping variables causing DIF is small, say one or two. However, if it turns out that a larger number of covariates have direct effects, it may be better to use a one-step LC model. In such a situation, a step-three analysis would require estimation of a large number of different D matrices (one for each DIF covariate combination), likely making the three-step approach less stable. Moreover, if say four out of the five covariates of interest show DIF, the step-one model will already contain four covariates, so why not include the fifth right away rather than proceeding with a three-step approach? Note that also in these types of situations, the first part of the proposed model-building strategy (in which one identifies the DIF covariates) remains very useful.

Vermunt (2010) and Bakk et al. (2013) showed that three-step LC analysis may not perform well when class separation is extremely low, that is, with entropy R-squared values below .5. On the other hand, the method was shown to be especially useful in moderate class separation conditions, where it performs very well. Though not investigated here, these results might also be expected to apply with the modified three-step LC model, when DIF is present.

Using a synthetic data set, we showed the new method reproduces the parameters perfectly in a theoretical *population*. However, we did not investigate bias in parameter and standard error estimates caused by *sampling* fluctuation. It can, however, be assumed that the results of earlier simulation studies apply here as well, though slightly larger sample sizes may be needed because the step-one model contains more parameters, and D matrices with more elements must be estimated. As shown by Bakk et al. (2014), step-three standard errors may be slightly underestimated, which can, however, be corrected. In Latent GOLD, this requires specifying the step-three model “manually,” that is, by providing the estimated D matrix in logit form as well as its estimated covariance matrix.

## Extensions

While this paper focused on LC models for categorical indicators, the proposed modified step-three approach can also be applied with other types of mixture models. It can be used in all situations in which covariates to be included in a step-three analysis have direct effects on the response variables. This includes latent profile models (Collier & Leite, 2017), mixtures of normal distributions (Gudicha & Vermunt, 2013), mixture growth models (Diallo & Lu, 2017), mixture regression models (Kim et al., 2016), and LC tree models (Van den Bergh & Vermunt, 2019). In addition, for three-step latent transition and latent Markov models with DIF, the proposed approach is also relevant (Asparouhov & Muthén, 2014; Di Mari et al., 2016).

While the focus of this paper was three-step LC analysis, exactly the same solution for dealing with DIF can be used in a two-step LC analysis (Bakk and Kuha, 2018). Step 1a, 1b, and 1c, which aim at finding the right step-one model, are also applicable with a two-step model. The step-two model is an LC model containing all covariates and all indicators, but with the measurement model parameters fixed to their step-one estimates. In the case of DIF, the measurement model will also contain the encountered direct effects. Actually, Di Mari and Bakk (2018) proposed an approach that is very similar to this one, with one important difference. They were not aware of the fact that the covariate-class associations should be included in the step-one model, because otherwise the direct effects will be overestimated.

## References

- Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using M plus. *Structural Equation Modeling: A Multidisciplinary Journal*, 21, 329–341. <https://doi.org/10.1080/10705511.2014.915181>
- Bakk, Z., & Kuha, J. (2018). Two-step estimation of models between latent classes and external variables. *Psychometrika*, 83, 871–892. <https://doi.org/10.1007/s11336-017-9592-7>
- Bakk, Z., Oberski, D. L., & Vermunt, J. K. (2014). Relating latent class assignments to external variables: Standard errors for correct inference. *Political Analysis*, 22, 520–540. <https://doi.org/10.1093/pan/mpu003>
- Bakk, Z., Tekle, F. B., & Vermunt, J. K. (2013). Estimating the association between latent class membership and external variables using bias adjusted three-step approaches. *Sociological Methodology*, 43, 272–311. <https://doi.org/10.1177/0081175012470644>
- Bakk, Z., & Vermunt, J. K. (2016). Robustness of stepwise latent class modeling with continuous distal outcomes. *Structural Equation Modeling: A Multidisciplinary Journal*, 23, 20–31. <https://doi.org/10.1080/10705511.2014.955104>
- Bandeen-Roche, K., Miglioretti, D. L., Zeger, S. L., & Rathouz, P. J. (1997). Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association*, 92, 1375–1386. <https://doi.org/10.1080/01621459.1997.10473658>
- Bolck, A., Croon, M., & Hagenaars, J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political Analysis*, 12, 3–27. <https://doi.org/10.1093/pan/mp001>
- Clogg, C. C., & Goodman, L. A. (1984). Latent structure analysis of a set of multidimensional contingency tables. *Journal of the American Statistical Association*, 79, 762–771. <https://doi.org/10.1080/01621459.1984.10477093>
- Collier, Z. K., & Leite, W. L. (2017). A comparison of three-step approaches for auxiliary variables in latent class and latent profile analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 24, 819–830. <https://doi.org/10.1080/10705511.2017.1365304>
- Dayton, C. M., & Macready, G. B. (1988). Concomitant-variable latent-class models. *Journal of the American Statistical Association*, 83 (401), 173–178. <https://doi.org/10.1080/01621459.1988.10478584>
- Di Mari, R., & Bakk, Z. (2018). Mostly harmless direct effects: A comparison of different latent Markov modeling approaches. *Structural Equation Modeling: A Multidisciplinary Journal*, 25, 467–483. <https://doi.org/10.1080/10705511.2017.1387860>
- Di Mari, R., Oberski, D. L., & Vermunt, J. K. (2016). Bias-adjusted three-step latent Markov modeling with covariates. *Structural Equation Modeling: A Multidisciplinary Journal*, 23, 649–660. <https://doi.org/10.1080/10705511.2016.1191015>
- Diallo, T. M. O., & Lu, H. (2017). On the application of the three-step approach to growth mixture models. *Structural Equation Modeling: A Multidisciplinary Journal*, 24, 714–732. <https://doi.org/10.1080/10705511.2017.1322516>
- Gudicha, D. W., & Vermunt, J. K. (2013). Mixture model clustering with covariates using adjusted three-step approaches. In B. Lausen, D. Van den Poel, & A. Ultsch (Eds.), *Algorithms from and for nature and life* (pp. pp. 87–94). Springer.
- Hagenaars, J. A. (1993). *Loglinear models with latent variables*. Sage.
- Janssen, J. H. M., van Laar, S., de Rooij, M. J., Kuha, J., & Bakk, Z. (2019). The detection and modeling of direct effects in latent class analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 26, 280–290. <https://doi.org/10.1080/10705511.2018.1541745>
- Kankaraš, M., Moors, G., & Vermunt, J. K. (2010). Testing for measurement invariance with latent class analysis. In E. Davidov, P. Schmidt, & J. Billiet (Eds.), *Cross-cultural analysis: Methods and applications* (pp. pp. 359–384). Routledge.
- Kim, M., Vermunt, J. K., Bakk, Z., Jaki, T., & Van Horn, M. L. (2016). Modeling predictors of latent classes in regression mixture models. *Structural Equation Modeling: A Multidisciplinary Journal*, 23, 601–614. <https://doi.org/10.1080/10705511.2016.1158655>
- Lee, S., Bulut, O., & Suh, Y. (2017). Multidimensional extension of multiple indicators multiple causes models to detect DIF. *Educational and Psychological Measurement*, 77, 545–569. <https://doi.org/10.1177/0013164416651116>
- Magidson, J., & Vermunt, J. K. (2001). Latent class factor and cluster models, bi-plots, and related graphical displays. *Sociological Methodology*, 31, 223–264. <https://doi.org/10.1111/0081-1750.00096>
- Masyn, K. E. (2017). Measurement invariance and differential item functioning in latent class analysis with stepwise multiple indicator multiple cause modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 24 (2), 180–197. <https://doi.org/10.1080/10705511.2016.1254049>
- McCutcheon, A. L. (1987). *Latent class analysis*. Sage.
- Muthén, L. K., & Muthén, B. O. (2015). *Mplus (Version 7.4)*. Muthén & Muthén.
- Nylund-Gibson, K., Grimm, R. P., & Masyn, K. E. (2019). Prediction from latent classes: A demonstration of different approaches to include distal outcomes in mixture models. *Structural Equation Modeling: A Multidisciplinary Journal*, 26, 967–985. <https://doi.org/10.1080/10705511.2019.1590146>
- Nylund-Gibson, K., & Masyn, K. (2016). Covariates and mixture modeling: Results of a simulation study exploring the impact of misspecified effects on class enumeration. *Structural Equation Modeling: A Multidisciplinary Journal*, 23, 782–797. <https://doi.org/10.1080/10705511.2016.1221313>
- Oberski, D. L., van Kollenburg, G. H., & Vermunt, J. K. (2013). A monte Carlo evaluation of three methods to detect local dependence in binary data latent class models. *Advances in Data Analysis and Classification*, 7, 267–279. <https://doi.org/10.1007/s11634-013-0146-2>
- Suh, Y., & Cho, S.-J. (2014). Chi-square difference tests for detecting functioning in a multidimensional IRT model: A Monte Carlo study. *Applied Psychological Measurement*, 38, 359–375. <https://doi.org/10.1177/0146621614523116>
- Van den Bergh, M., & Vermunt, J. K. (2019). Latent class trees with the three-step approach. *Structural Equation Modeling: A Multidisciplinary Journal*, 26, 481–492. <https://doi.org/10.1080/10705511.2018.1550364>
- Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis*, 18, 450–469. <https://doi.org/10.1093/pan/mpq025>

- Vermunt, J. K., & Magidson, J. (2016). *Technical guide for Latent GOLD 5.1: Basic, advanced, and syntax*. Statistical Innovations Inc.
- Vermunt, J. K., & Magidson, J. (2020). *Upgrade manual for Latent GOLD 6.0*. Statistical Innovations Inc.
- Woods, C. M. (2009). Evaluation of MIMIC-model methods for DIF testing with comparison to two-group analysis. *Multivariate Behavioral Research*, 44, 1–27. <https://doi.org/10.1080/00273170.802620121>

## Appendix A1. Implementation in Latent GOLD 6.0 software

Latent GOLD 6.0 (Vermunt & Magidson, 2020) implements the method for dealing with DIF discussed in this paper. The covariate case illustrated with the data from Hagenaaers (1993) can be dealt with either using the Cluster, DFactor, and Step3 point-and-click modules or the Syntax module. Step-three distal outcome models with DIF should be estimated with Syntax.

### Using Latent GOLD point-and-click modules

The step-one analysis and step-two classification can be performed with the Cluster or DFactor module. Step 1a, in which one determines the number of classes, is done using only Indicators. Step 1b, in which one determines which external variables cause DIF, involves including them, one at a time, as a covariate in the model. The Bivariate Residuals output can be used to determine whether there are large residual associations caused by direct effects. These direct effects can be included in the model by checking the appropriate effect on the Residuals tab. Step 1 c involves including all DIF Covariates simultaneously and checking the appropriate direct effects on the Residuals tab. The step-two classification to an output data file based on the Step 1 c model can be performed with the options on the ClassPred tab; that is, by checking Classification-Posterior, specifying the name of the newly created data file, and selecting the other variables that should be written to the output data file.

After opening this output data file in Latent GOLD, one can activate the Step3 module. The posterior probabilities *clu#1*, *clu#2*, etc., are selected in the Posteriors box, and the covariates to be used are moved to the Covariates box. For each covariate, one should specify the scale type (Nominal, Numeric, B-spline), as well as whether it is a DIF covariate. By default, the “proportional” assignment and “ml” adjustment are used, which can be changed to “modal” and “BCH”, respectively. On the Technical tab, one may modify the missing values setting and/or other technical settings, and on the Model tab, one may include interaction terms.

### Using Latent GOLD Syntax module

While steps 1a, 1b, 1 c, and 2 can almost always be done using the point-and-click modules, some users may prefer using the Syntax. Below is example Syntax for an LC model with six categorical indicators (*y1*-*y6*) and a single covariate (*z1*) with direct effects on two of the indicators (as in our synthetic example), which also includes classification to an output data file:

```
options
  output parameters standarderrors profile bivariateresiduals
  estimatedvalues;
outfile 'classification.txt' classification keep = z2;
variables
  dependent y1, y2, y3, y4, y5, y6;
  independent z1 nominal;
  latent Class nominal 3;
equations
  Class <- 1 + z1;
  y1 <- 1 + Class;
  y2 <- 1 + Class;
  y3 <- 1 + Class + z1;
  y4 <- 1 + Class;
  y5 <- 1 + Class;
  y6 <- 1 + Class + z1;
```

The following is example Syntax for a step-three model with two covariates (*z1* and *z2*) affecting class membership, where *z1* is a DIF covariate:

```
options
  step3 modal ml;
  output parameters standarderrors = robust profile = posterior
  probmeans = posterior estimatedvalues;
variables
  independent z1 nominal;
  latent Class nominal posterior = (Class#1 Class#2 Class#3) dif = z1;
equations
  Class <- 1 + z1 + z2;
```



If z2 is a distal outcome (say a continuous variable) and z1 a DIF control variable, the step-three model specification would be as follows:

```
options
  step3 modal bch;
  output parameters standarderrors = robust profile probmeans = posterior
    estimatedvalues;
variables
  independent z1 nomina;
  dependent z2 continuous;
  latent Class nominal posterior = (Class#1 Class#2 Class#3) dif = z1;
equations
  Class <- 1 + z1;
  z2 <- 1 + z1 + Class;
```

Note that we changed the step3 correction method from ml to bch, which is the preferred method for continuous distal outcomes.