ABSTRACT

An Examination of Parameter Recovery in Latent Transition Models with Distal Outcomes

Courtney A. Moore, M.A.

Chairperson: Grant B. Morgan, Ph.D.

Latent transition analysis (LTA) is an increasingly popular research method used to categorize subsets of individuals within a population. The current study is purposed to investigate parameter recovery of a distal outcome effect in an LTA model. All models have a two-class solution with two time points. Design factors of interest include sample size, class prevalences at Time 1, the transition parameter, and the distal outcome effect size. ANOVA was used to seek out any practically significant effects ($\eta^2_P \geq 0.14$) of these design factors on the raw bias (RB) of the distal outcome estimate. The results revealed no practically significant effects, meaning that the LTA model accurately estimated the distal outcome effect size under all specified conditions. Future research could expand upon this study by including different numbers of classes, indicators, time points, distal outcomes, and other auxiliary variables in addition to the distal outcome(s).

An Examination of Parameter Recovery in Latent Transition Models with Distal Outcomes

by

Courtney A. Moore, B.A.

A Thesis
Approved by the Department of Educational Psychology

Susan Johnsen, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of

Master of Arts

Approved by the Thesis Committee

Grant B. Morgan, Ph.D., Chairperson

Wade C. Rowatt, Ph.D.

Terrill F. Saxon, Ph.D.

Accepted by the Graduate School
August 2017

J. Larry Lyon, Ph.D., Dean

Page bearing signatures is kept on file in the Graduate School.

Copyright © 2017 by Courtney A. Moore

All rights reserved

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

Introduction

Classification is the act of evaluating sets of individuals and grouping them per some characteristic or array of characteristics. The general idea behind classification involves individuals in each group being similar to other individuals within that group and dissimilar to those in other groups in regards to a specified characteristic of interest. Compared to other means of analysis that assume a single continuum on which all data points lie, this method allows the researcher to evaluate the theory that there may be more than one underlying category or trend within a data set.

Two benefits of classification are simplification and interpretability. Say a data set includes 5000 individual cases. Analytically categorizing these individuals could result in perhaps three classes that a researcher could treat as three single entities. Interpretability is simultaneously aided by reducing a large number of individual cases to these three classes of individuals. The researcher can now examine the shared features within each of the three classes to determine that which makes the individuals in each class similar to one another and different from those in the other two classes. From this, the researcher can draw the conclusion that in regards to the characteristic(s) being studied, there are three distinct types of people or response patterns.

As previously mentioned, more traditional modes of data analysis assume a single continuum that inhabits every data point within a sample. Thus, inferences are based on the placement of every included individual within this single distribution. While these analyses are appropriate for theoretically supported hypotheses that also assume a single distribution or continuum of people for a certain characteristic, other theories support the possibility of distinct subsets of people within the greater population. Classification allows attention to be directed to these analytically specified subsets of individuals, and conclusions are drawn based on similarities within and differences between the groups.

Within education, categorizing students enables the development of specialized interventions or teaching methods to fit the shared needs of subgroups of students. While creating individualized curricula or educational programs for each of the 50.4 million students in the public education system would be ideal, doing so is simply unrealistic based on funding and personnel. Therefore, identifying sets of students with similar characteristics and needs makes possible the efficient creation of interventions with the propensity to affect a greater number of students with that need. For example, say students are categorized by their specific learning disabilities. Programs and strategies could then be created for and tested within each group based on the common learning disability within that subset of students. Compared to attempting to address each student's needs individually and subsequently create an individualized intervention, the classification process would enable more students to be affected at one time. This is not to say that there are no students for whom individualized support is merited. This example is simply meant to highlight the efficiency and practicality of the classification process in educational settings.

Classification analyses can also be performed at two (or more) time points, adding a longitudinal element to the process. Such an analysis is purposed to investigate not only which classes individuals are categorized into, but also to see if individuals transition into other classes over time. Evaluating these transition patterns or probabilities can give researchers a sense of not only how the specified characteristic(s) manifests itself in the population, but how the trend either changes or maintains consistency over time. As an example, say first-time teachers of third-graders are classified during their first month according to the teaching methods they practice. Then, they are classified again at the end of the school year. A researcher can then evaluate the classes of teachers as well as the transition probabilities over time to see both which methods are being utilized and which methods are switched to or from over the course of the year.

Yet another layer that can be added to the classification process includes assessing the effect of class membership on an outcome, known as a distal outcome. For example, say classification is based on high school students socioeconomic status (SES), grades in core classes (math, science, history, English), and motivation. The class to which one belongs could then be related to that persons salary 10 years after graduation. Perhaps a class of students with low SES, moderate grades, and high motivation would be related to obtaining a higher salary than those in a class characterized by high SES, low grades, and low motivation. In this example, salary is the distal outcome.

The addition of the distal outcome variable to the classification analysis has the potential to provide increased information about the importance of class membership. With classification alone, one can examine subgroups of people within a population. Studying classes over time allows a researcher to investigate the patterns of transition between or among these subgroups, and adding the distal outcome variable can give insight into the possible predictive implications of class membership.

Purpose of Study

The amount of people in each of the classes in a classification analysis is stated as a proportion. These proportions known as prevalences have the ability to vary, as do the transition patterns between classes over time. The affect that class prevalences, transition probabilities, and other design factors have on the ability of an LTA model to correctly predict a distal outcome has yet to have been studied. Without methodological research in this area to say otherwise, there is the possibility that the analysis might not be trustworthy in consistently producing accurate results. For this reason, the current study addresses this need.

A Monte Carlo simulation was performed in which specifications of the classification model (e.g. class prevalences, transition probabilities, sample size, number of classes, time points, effect size associated with the distal outcome) were set to reflect the findings reported in recent applied research. The data sets produced from these specifications were then analyzed to test the performance of the procedure, with the outcome measure being raw bias (RB), which is the difference between the estimated effect size of the distal outcome and the true effect size. Greater RB values represent a greater difference between estimated and specified effect sizes and thus a less accurate prediction mechanism.

CHAPTER TWO

Literature Review

As was previously established, classification is often utilized in applied research when the population is hypothesized to be comprised of two or more characteristically distinct subgroups of individuals. Identifying these subgroups thus allows the researcher to delve into the meaningfulness of these distinctions between the groups as well as the similarities accounted for within each of the groups. Within education, methods used for classification may provide increased knowledge regarding the underlying composition of or trends within student populations, which can in turn direct the creation of interventions, curricula, and teaching methods that align with the patterns identified via classification.

Latent Class Cluster Analysis

Latent class cluster (LCC) analysis is a commonly used, model-based method of classification analysis applied to multivariate data sets (P. Lazarsfeld & Henry, 1968). This family of procedures is often considered within the framework of structural equation modeling (SEM) and have been identified by various names, such as model-based clustering (Banfield & Raftery, 1993), mixture likelihood approach to clustering (Everitt, 1993; McLachlan & Basford, 1988), and finite mixture modeling (McLachlan & Peel, 2000). The term finite mixture modeling refers to the hypothesis within LCC analysis that an analytic data set is a mixture of a finite number of characteristically differentiable and mutually exclusive classes. For this reason, LCC should only be applied when the hypothesis supports the presence of multiple classes within the data. Even though a multi-class hypothesis is necessary to justify the use of LCC, both the form and frequency of the classes are unknown prior to analysis. Form involves all parameters used to differentiate the classes (e.g. means, variances) (Vermunt & Magidson, 2002), and frequency refers to the number of classes within the data set. Although this is true of all classification methods, one distinguishing characteristic of LCC is the theorized presence of a categorical latent variable (Collins & Lanza, 2010). A latent variable is one that cannot be directly observed and must therefore be approximated through the use of measurable indicators of that latent construct. The categorical nature of the latent variable is supported by the presence of multiple classes or categories of individuals within the target population of a classification analysis. Although latent variables are treated as errorless constructs, the measured variables used to estimate them are simply indicators of the construct and not perfect representations of the latent variable. Therefore, measurement error is present in this approximation.

Assumptions

As mentioned previously, a characteristic of all LCC models is the presence of a latent variable (Heinen, 1996). This latent variable is further theorized to be wholly responsible for the connection between the observed variables of the data set, a concept known as local, or conditional, independence (Clogg, 1995; McCutcheon, 1987; Vermunt & Magidson, 2002). Local independence warrants that the observed variables though correlated before analysis should be uncorrelated with one another after accounting for the latent variable. This occurs when the latent classification variable fully explains the relationships within the data set. When this assumption is tenable, ones response to an indicator variable is dictated, or caused, by the latent variable.

The assumption of independence is considered local because it is theorized to be true within each class (Collins & Lanza, 2010), with each class representing a distinct, homogenous population within the data set. Because these separate populations within the data are hypothesized to be homogenous, each individual within a class has the same latent probability distribution as other members of the class. Additionally, class membership is considered to be mutually exclusive as well as exhaustive, as each case within a data set is assumed to belong to one class and one class only (Clogg, 1995; Heinen, 1996).

Benefits

LCC allows for the existence of multiple subpopulations within a data set (P. Lazarsfeld & Henry, 1968). This characteristic enables researchers to address hypotheses about these subpopulations beyond traditional analyses that assume a single underlying population (Collins & Lanza, 2010). Once each subpopulation or class is statistically distinguished, each case is placed into the class to which it best belongs based on its similarity to and differences from other cases. Even though each case is placed into one class only, LCC acknowledges the potential for misclassification by estimating the probability that each case belongs to each of K classes (Asparouhov & Muthén, 2014a; Bakk & Vermunt, 2016). These probabilities are based on how well the distribution of indicators for each case aligns with the distribution for each class. Lastly, model-based approaches like LCC provide researchers with the ability to free and fix model parameters as best serves the purposes and hypotheses of the research. Freed parameters are estimated during analysis, while others that are not of interest can be fixed to a certain value or even fixed to be equal with other parameters. The choice to free and fix model parameters should be grounded in previous research and theory.

Limitations

As previously mentioned, the data set must include multiple populations (i.e. classes) in order to warrant the use of LCC analysis. Moreover, each class has to be sufficiently large and distinguishable from the other classes in order for LCC to unveil the underlying class structure of a data set. Because of this limitation, LCC might not detect rare classes, especially when the total sample size is small and class enumeration is high (Thompson, 2007). It is also assumed that the model under analysis is correctly specified, which is an assumption in any analysis involving a model although it is often unstated. Even if a model is correct, though, it is virtually impossible for that model to include every source of variability regarding the latent variable. All models are thus considered overly restrictive, meaning that the goal, then, is to have the least restrictive model possible while maintaining interpretability and parsimony. Lastly, the independence assumption may not always be realistic in application. Theoretically, there could be sources of variability outside of the latent variable, possibly causing the observed variables to remain related even after accounting for the latent variable.

Although LCC has its limitations, many of these concerns can be buffered when the research design is grounded in previous research and theory. Several of these issues can be positively aided through the collection of a larger sample as well as with the use of theoretically-based indicator variables that really do approximate the latent variable. These measures should also be known to result in reliable decision-making and valid interpretation(s) of data. Concerning local independence, there is the option within LCC to loosen the independence assumption should ones theory support the decision to do so. As a result, however, the researchers interpretation of the classes may be more challenging.

Latent Class Analysis

Relatively recent technological and statistical advances have allowed for increased accessibility and efficiency in utilizing LCC methods, thus contributing to an increase in its use within applied research settings (Vermunt & Magidson, 2002). Latent class analysis (LCA) is a widely-used submodel of LCC and the focus of the current study. Developed by P. F. Lazarsfeld (1950), LCA employs categorical (i.e., nominal or ordinal) indicator variables to approximate the latent construct. Like all LCC models, LCA is applied when the population of interest is hypothesized to be comprised of two or more subgroups, or classes, of people. Should this theory be supported (and if the data are collected and analyzed using best practices), LCA may produce solution options that analytically separate groups according to similar and dissimilar traits. A path diagram depicting an LCA model with 5 indicators is shown in Figure 1. As shown in Figure 1 with the direction of the arrows, latent class membership C_1 is responsible for the values of indicators $U_1 – U_5$. In other words, each study participants responses to items $U_1 – U_5$ are determined, with some measurement error, by the latent class to which she or he belongs.

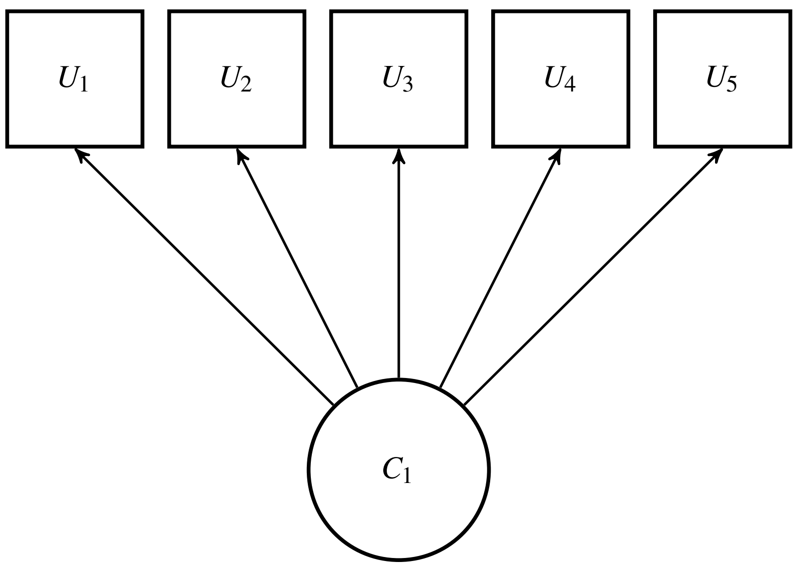

Figure 2.1. Path diagram of latent class analysis model with five indicators.

LCAs use of categorical indicator variables is that which distinguishes this method from other LCC methods. This property of LCA allows for the evaluation of response patterns among individuals, which is why this analysis is often used with item-level data (Collins & Lanza, 2010). This is largely evident within education, as selected-response items (e.g., multiple choice items) are commonly used on educational assessments. In a basic sense, LCA uses each persons response pattern to analyze how similar and dissimilar people are from one another. These similar and dissimilar patterns are then grouped accordingly into classes.

As an example, suppose that a test included four items, each of which was multiple choice and coded as either correct (1) or incorrect (0). Test-takers response patterns are then a four-digit combination of zeros and ones that reflect their test performance. Thus, it follows that a student who answered every question correctly would have a 1111 response pattern. Perhaps another student correctly responded to the first two items but incorrectly to the last two, making her response pattern 1100. From an evaluation of all response patterns within the sample, the frequency of individuals with each response pattern can be determined (Heinen, 1996). The number of possible response patterns is determined by the number of items along with the number of response options for each item. In this example, there are four items, each with two response options (correct or incorrect). Therefore, there are 24 = 16 possible response patterns. This means that the contingency table for this example will include 16 cells to accommodate the 16 unique response patterns, and each cell will represent the number of individuals with that specific response pattern (Collins & Lanza, 2010). The table can then be used to identify any response patterns that are more or less common than others. This evaluation can provide the researcher with an idea as to the total distribution of individuals according to their performance on the four-item test.

The contingency table of response patterns is then utilized to perform LCA, which is designed to (1) identify the set of classes that would produce these response patterns, (2) evaluate the prevalence of each class, and (3) estimate parameters (Φ) for the model (Collins & Lanza, 2010). The LCA model returns the probability of observing a certain set of responses based on the probability of membership within each latent class. The LCA model differentiates itself from other LCC models in that it is adapted for categorical data. This means that it is unnecessary to uphold strict assumptions about the distributions of the included variables as they are categorical. Further, if all items are dichotomous, the joint distribution can be treated as binomial (Collins & Lanza, 2010; Everitt, 1993).

The general LCA model is expressed as:

$$f(y_i) = \Sigma_{k=1}^{K} \pi_k f_k(y_i|\rho_k, \Sigma_k) \tag{2.1}$$

in which the distribution of the indicators (y_i) relies on the probability that a specific response pattern is observed for case π_i. The parameters estimated within the equation include the group mixing weight π_k, and the values within that classs variance-covariance matrix, Σ_k. Another is the probability ρ_k of an individual observed response on an indicator variable based on class membership, the values of which can range from 0 to 1 (Collins & Lanza, 2010).

Latent Transition Analysis

Latent transition analysis (LTA) is the application of LCA to longitudinal data. LTA can provide insight into not only class prevalences at each time point, but also the patterns of transition among classes between subsequent time points (Collins & Lanza, 2010). In other words, this method can be used to evaluate both if and in what capacity individuals move from one class to another over a certain amount of time. Yet another way to conceptualize LTA is as a method of testing the stability of an LCA solution over a period of time (Morgan, Wells, Andretta, & McKay, 2016). These patterns of change between time points are quantified in transition probabilities, which are given for every possible transition pattern. For example, in a two-class solution, probabilities are established for transitions from Class 1 to Class 2, Class 2 to Class 1, and also for the percentage of people who maintain their class membership (Class 1 to Class 1, Class 2 to Class 2).

A general LTA model with two classes at two time points can be expressed:

$$\tau_{ikm} = P(C_{i2} = k|C_{i1} = m) = \frac{exp(\alpha_2 + \gamma I(C_1))}{exp(\alpha_2 + \gamma I(C_1)) + 1'} \tag{2.2}$$

where τ_{ikm} is the probability of belonging to class k at time 2 conditional on class membership at time 1, γ is the predictive relationship of class 2 from class 1, and I(C_1) is simply a dummy-coded indicator of belonging to class 1 at time 1. The path diagram of this model is shown in Figure 2 with five indicators per time point. The model can of course be extended to multiple classes per time point and more time points, but the two-class, two-time point model is shown here because it is the focus of this study (see Chapter Three for more detail).

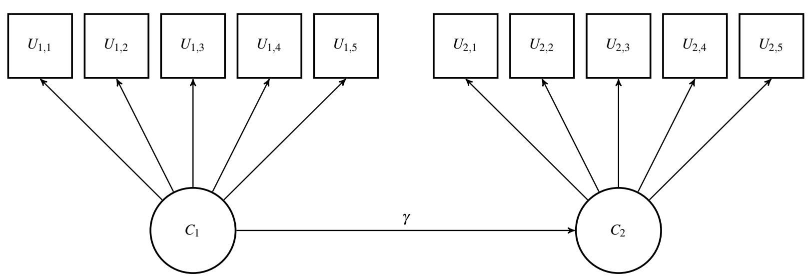

Figure 2.2: Path diagram of latent transition analysis model with two time points and five indicators at each time point.

Distal Outcomes

When added to any model, distal outcomes offer a predictive element. In regards to LTA, the class prevalences and transition probabilities could be related to a future outcome. Thus, the distal outcome is theorized to be predicted by class membership and/or transition pattern over time (Lanza, Tan, & Bray, 2013; K. L. Nylund, Muthén, Nishina, Bellmore, & Graham, 2006). In other words, membership in one class could predict a certain level of a future outcome for the individuals within that class, whilst another class is predicted to be at a different level regarding the distal outcome. A path diagram showing a distal outcome is shown in Figure 3. As can be seen, the model simply extends the LTA model to include the temporally distal outcome, X.

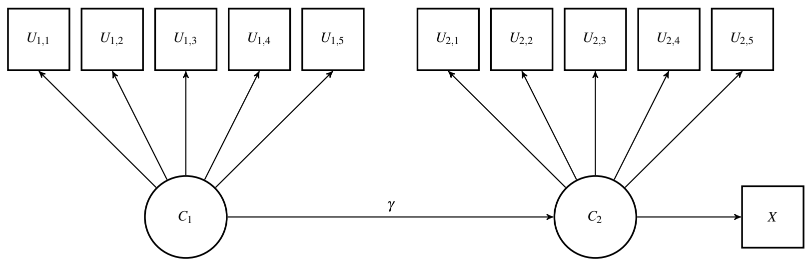

Figure 2.3: Path diagram of latent transition model with two time points, five indicators per time point, and one distal outcome.

One- Versus Three-Step Approaches

The addition of distal outcomes in latent variable models has brought about methodological research to investigate whether a one-step or a stepwise analytical approach should be used. The one-step approach estimates the measurement model and the regression model simultaneously. Since the estimations simultaneously occur, the one-step method essentially treats the distal outcome which is assumed to be normally distributed as another indicator of the latent variable. The attractiveness of the one-step approach comes in that it is more efficient than stepwise approaches by estimating all parameters at once rather than breaking the process into separate steps. However, this method does not agree with the theory that the distal outcome is predicted by the latent variable rather than the other way around. Additionally, the distal outcome must be held to stricter assumptions since it is being analytically treated as an indicator with this method.

Stepwise analytical approaches of which four will be discussed respect the directionality of the relationship between the latent variable and the distal. The measurement model is estimated prior to the regression model, with the estimated parameters from each step being utilized in the subsequent steps of analysis. Though stepwise analyses are in line with the theory that the distal outcome is predicted, these methods are not without their criticisms.

Modal Class Approach

The modal class approach assigns each individual to a class directly following the estimation of the measurement model. These assignments are based on ones posterior probability of being in each of the classes, which assumes that ones class membership is certain rather than probabilistic. This can lead to biased parameter estimates and high classification error, subsequently leading to biased estimates in further steps of the analysis (i.e. estimation of distal effect size).

Lanza-Tan-Bray Approach

The Lanza-Tan-Bray (LTB) method begins by treating the distal outcome as an indicator of the latent variable in order to estimate distal outcome means for each of the classes (Lanza et al., 2013). It then reverses the direction of estimation using Bayes theorem to estimate the distal outcome from the latent variable, either through the use of kernel densities or the sampling distribution of the distal outcome (Asparouhov & Muthén, 2014a). The LTB method is purposed to maintain the latent class solution arrived at in the first step (Lanza et al., 2013). Issues with this approach include that the estimation of the distal outcome is biased when the distal has outliers (Bakk, Oberski, & Vermunt, 2014) and the model cannot include covariates (Asparouhov & Muthén, 2014a). Other limitations and information can be found in these articles as well as in Bakk and Vermunt (2016).

Modified BCH Approach

Adapted from the original BCH approach (Bolck, Croon, & Hagenaars, 2004), the modified BCH approach (Vermunt, 2010) was extended by Bakk and colleagues (2013) to include distal outcomes. Similar to the LTB approach, the modified BCH approach retains the latent variable model estimates obtained in the first step throughout the analytical process. It differs, however, in that its performance does not rely on as strict of assumptions regarding the distal outcome (Asparouhov & Muthén, 2014b; Bakk & Vermunt, 2016). Nevertheless, estimation biases may occur in models with categorical distals (in some instances) or when there are more than one distal outcome (Asparouhov & Muthén, 2014b; Bakk, Tekle, & Vermunt, 2013). As is consistently found in methodological research, small sample size along with poor class separation can lead to an underestimation of standard errors (Bakk & Vermunt, 2016).

Three-Step Approach

Inspired by the modified BCH method, the three-step approach first performs LCA, excluding distal outcomes and any covariates in the model (Asparouhov & Muthén, 2014a). From this step, the most probable class assignments as well as the posterior probabilities of being in each class are computed. The second step of the process involves analyzing the misclassification error in the first step of analyses. These estimates are then utilized in the following step, as is also the case in the modified BCH method. Step three applies a specific log-likelihood function to obtain estimates of the distal outcome, which is assumed to be normally distributed within each class. It is also in this final step that the assumption of measurement invariance takes hold in that the class structure at Time 1 (decided upon in the LCA during step one) is maintained at each subsequent time point. People can transition between classes over time, but the structure of the classes does not vary over time. While this approach takes into account misclassification error, it does underestimate standard errors for which a correction was developed (Bakk et al., 2014).

Conclusion

The debate of whether to use the one- or three-step approach is ever present among methodological researchers. While the one-step is preferred for efficiently estimating all pieces of the model simultaneously and research has found little difference between the performance of the separate methods theoretical expectation tends to be most accurately reflected in the three-step method. The distal outcome is believed to be predicted by latent class membership. Therefore, since it is temporally distant from the other model parameters, proponents of the three-step method argue that it should be estimated after the rest of model (Nylund-Gibson, Grimm, Quirk, & Furlong, 2014). In lieu of this theoretical argument, the three-step model will be utilized in this study.

Examples of LTA in Applied Research

LTA is becoming an increasingly popular analysis in social, psychological, and educational research settings, to name a few. As noted in Chapter One, classification methods are often used in educational research in the pursuit of identifying common trends among students and teachers that may be utilized in the creation of interventions, teaching methods, and even curricula. LTA can provide information about underlying trends and subgroups within student populations, as is evidenced by Swanson, Kudo, and Guzman-Orths (2016) study of the risk of reading disability in elementary students. The sample included 489 early elementary students for whom English was their second language, Spanish being their first. Four latent classes emerged at each time period of the LTA (Year 1, Year 2, Year 3), with the classes being labeled nonbalanced bilinguals-good readers, balanced bilinguals-good readers, reading disabled, and classroom inattentive children (Swanson, Orosco, & Lussier, 2012). LTA results showed that 10% of students in the inattentive children class transitioned to the reading disabled class by Year 3, while the rest of the classes remained relatively stable over time. The findings support the overall stability of reading disability over time, while it was also discovered that cognitive skills rather than oral and phonological abilities predict membership in the reading disabled class.

While much research within education focuses on learning, another portion focuses on the social and developmental aspects of educational settings. K. L. Nylund et al. (2006) used LTA to study victimization in 1,300 public middle school students at three time points. Their model included covariates, first and second order effects, mover-stayer variables, and distal outcomes. The three classes discovered were labeled victimized, sometimes victimized, and nonvictimized (K. L. Nylund et al., 2006). The researchers found that students were more likely to transition to a less victimized group rather than to become more victimized over time. Additionally, those students who were victimized were less likely to feel safe in school, as well as more likely to be socially anxious and depressed. Once in high school, those who were victimized throughout middle school reported more physical health issues and social anxiety. LTA with model extensions such as the inclusion of distal outcomes allowed the researchers to identify not only patterns of victimization over time, but also the widespread and long-lasting effects with which its associated.

McNeil Smith and Fincham (2016) analyzed racial discrimination patterns among 711 African American eighth- and eleventh-graders. LTA revealed three classes of individuals based on both the amount of racial discrimination they experience and from whom the discrimination is received. The first class was the Teacher-Peer Perceived Racial Discrimination (PRD), which represented students who had a high probability of being racially discriminated against by both teachers and peers. The second group was labeled Teacher PRD and include those who experienced discrimination from teachers but not peers. The final class was No PRD, which included students who reported low levels of racial discrimination from both teachers and peers. The LTA model included covariates of gender and family income, and analysis revealed that males were far more likely than females to belong to the Teacher-Peer and Teacher PRD groups. Students were also more likely to report lower levels of racial discrimination at the second time point as shown by the transition patterns. In this study, LTA revealed the existence of distinct discrimination experiences among African American students while also providing probabilities of moving between these distinct groups over time.

Previous Methodological LTA Research

The increased use of LTA in applied research is a relatively recent development, which is additionally reflected in the methodological research that has been done in this area. While the longitudinal piece of LTA requires inspection beyond what can be known from LCA research, much methodological research of LCA can be applied to LTA. For example, Bakk and Vermunt (2016) compared four bias-adjusted stepwise methods (maximum likelihood (ML) with equal variances, ML with unequal variances, BCH, and LTB) on their estimation performance in LCA with distal outcomes. When assumptions were met, all methods produced unbiased estimates of the distal outcomes. However, only the BCH approach could do so when assumptions were violated. These stepwise methods are applied to both LCA and LTA model estimations (as discussed previously), which makes the findings from this study important in their implications for LTA.

Clark and Muthén (2009) investigated the effect of data and LCA model characteristics (e.g. sample size, entropy) on the estimates of covariates. The main focus was on the comparisons of five different regression types (most likely class membership regression, probability regression, probability-weighted regression, pseudo-class regression, and single-step regression) in the estimation of covariate effects and other parameter estimates. These comparisons analyzed differences in coverage, or the proportion of simulated data sets that included the previously established parameter values (mean-square error, standard error, covariate effects). They found that these regression methods produce adequately accurate estimates as long as either entropy is high or the true covariate effect is zero. Under alternative conditions, such as lower entropy or a covariate effect of .5, they discovered that the single-step method provided the most accurate estimate of the covariate. However, the researchers do note that having more than one auxiliary variable in the model could be problematic.

Asparouhov and Muthén (2014a) studied performance of the one-step method and several three-step methods with both LCA and LTA with auxiliary variables. A distal outcome was included in their analysis of LCA and comparisons were drawn between the three-step and LTB methods. Both categorical and continuous distal outcomes were examined, the categorical with which only LTB was applied. They found that when the distal outcome was not normally distributed as assumed, both the one- and three-step methods could lead to changes in the latent class estimates. They also found that these two methods were less trustworthy when the latent class variable is affected by the inclusion of the distal outcome. Conversely, relatively accurate estimations were still produced by both the LTB and pseudo-class (PC) method (Rubin, 1987) in the latter scenario. With LTA, the study included one simple model and one model with covariates and measurement invariance, both of which were analyzed using the three-step method. The researchers concluded that the three-step method produces dependable estimates for both models, and that this conclusion can be extended to LTA models with more than two latent classes and covariates that are solely included in the third step of analysis.

Turning to LTA research, Cho, Cohen, Kim, and Bottge (2010) combined LTA with a mixture Rasch model (MRM). One described benefit of the LTA-MRM is that such a model allows within-class variability while LTA on its own assumes within-class homogeny. The model can thus be used to analyze varied effects within classes rather than solely the differences between them. For this study, the estimated parameters included item difficulty, ability, transition patterns, and class membership. As is common with most classification methods, the LTA-MRM did not perform well with small sample sizes. As is additionally common with item analysis methods, the model also performed best when the evaluated measure contained a greater number of items. This approach to mixture modeling provides an innovative option for real-world data analysis in which one seeks to simultaneously answer questions about classes of individuals and items through the use of LTA.

Purpose of Current Study

As can be seen, there exists methodological research on LCA with distal outcomes, as well as that of LTA models with other auxiliary variables such as covariates. However, there has yet to be methodological research on LTA models with distal outcomes and the effects of model parameters on distal outcome effect estimation. Distal outcomes are being increasingly utilized in longitudinal research using classification analyses, which is why the current study will investigate potential influences on a distal outcome in LTA.

CHAPTER THREE

Methods

The current study looks at a longitudinal latent class model with dichotomous indicators and a distal outcome. This type of model is known as a latent transition analysis (LTA) model with distal outcome. As such, a latent class model is fitted at multiple time points and the transition between classes across time is one outcome of interest. When distal outcomes are also included in the model, the effect of transitions between classes across time can also be examined with respect to a temporally distant (i.e., distal) outcome. The generalized latent class model, which is estimated at each time point, can be expressed as:

$$f(\boldsymbol{y_i}|\boldsymbol{\Phi}) = \Sigma_{k=1}^{K} \pi_k \Pi_{j=1}^{J} f_k(\boldsymbol{y_i}|\boldsymbol{\theta_k}) \quad (3.1)$$

where $\boldsymbol{y_i}$ represents the set of scores for case i across variables, K is the number of underlying classes, J is number of indicators, π_k represents the prior probability of being in class k (i.e., class prevalence), and $\boldsymbol{\theta_{jk}}$ is the set of model parameters (Vermunt & Magidson, 2002). A more thorough overview of LCC is included in Chapter Two. As is mentioned in that description, local independence is an assumption of LCC models (Clogg, 1995; McCutcheon, 1987; Vermunt & Magidson, 2002), warranting the specification of a univariate distribution function for each yij. Class parameters are estimated using k binomial distributions, which are binomial since the indicators are dichotomous.

The longitudinal, or latent transition, model with two classes at each of two time points can be expressed:

$$\tau_{ikm} = P(C_{i2} = k|C_{i1} = m) = \frac{exp(\alpha_2 + \gamma I(C_1))}{exp(\alpha_2 + \gamma I(C_1)) + 1'} \quad (3.2)$$

where τ_{ikm} is the probability of belonging to class k at time 2 conditional on class membership at time 1, γ is the predictive relationship of class 2 from class 1, and $\mathrm{I}(C_1)$ is simply a dummy-coded indicator of belonging to class 1 at time 1. The class structure is assumed to be invariant across time.

Simulation and Estimation

Mplus (L. K. Muthén & Muthén, 2017) was used in the current study for simulation and estimation because it includes advanced options within Monte Carlo simulation (B. Muthén, 2003). Estimation was automated via the MplusAutomation package (Hallquist, 2011) in R (R Development Core Team, 2017) because the package can write and run multiple Mplus command files with a single R command line.

In estimating parameters in an LCC/LTA model, the number of classes must be determined at each time point from which data were collected (Hunt & Jorgensen, 2003). Classifying cases into groups is then done through an estimation method such as maximum likelihood (ML). ML uses an individuals set of scores across indicators to provide the likelihood that that individual belongs to class k based on the similarity between an individuals response profile and the characteristics of each class. This likelihood is achieved via an alternative form of the expectation-maximization algorithm, which functions to maximize the likelihood function for the K cluster model with specified class enumeration. The achieved parameter estimates are those that most likely produced the data set, as they maximize the likelihood function.

One assumption of ML is multivariate normality across indicators, which is rarely achievable with dichotomous data due to constrained variances. Robust maximum likelihood is then used rather than ordinary ML for LCC models with dichotomous indicators (L. K. Muthén & Muthén, 2010). The standard errors of parameter estimates produced in robust maximum likelihood estimation are robust to a violation of the assumption of multivariate normality (L. K. Muthén & Muthén, 2010).

Since dichotomous indicators are present in the current study, robust maximum likelihood estimation will be used over ML. For these dichotomous indicators, class-specific proportions (i.e., response probabilities) will be estimated.

Although estimation is purposed to maximize the likelihood function, these methods may identify local maxima rather than the global maximum (K. Nylund, Bellmore, Nishina, & Graham, 2007). This issue is combated by having multiple sets of random starting values for the estimation algorithm. The global maximum is identified when these sets converge on the same estimate. In line with recommendations from previous research and in order to maximize the probability of arriving at the global rather than a local maximum the number of random start sets will be increased to 1000 with 20 optimizations for each set (K. Nylund et al., 2007) from the default of 20 sets of starting values and four optimizations in the final stage.

Steps for Estimating LTA Model with Distal Outcome

For the current study, estimation will follow a specified sequence in line with the modified BCH estimation procedure (Bakk & Vermunt, 2016), which was chosen above other three-step approaches for taking into account misclassification error. This procedure is carried using the following steps:

(1) An LCA model is fitted to the data that constrains the latent class parameters to be equal across time points (i.e., measurement invariance). The parameters estimates are saved as starting values for subsequent analyses. The latent transition parameter is NOT estimated in this step.

(2) An LCA model is estimated for time 1 using the starting values from Step 1.

(3) The classification probabilities for each individual from Step 2 are saved.

(4) An LCA model is estimated for time 2 using the starting values from Step 1.

(5) The classification probabilities for each individual from Step 4 are saved.

(6) The LTA model is estimated and uses the classification logits from Steps 3 and 5, which accounts for the degree of uncertainty with which each individual was assigned to each class at each time point. The LTA model involves regressing class membership at Time 2 onto class membership at Time 1. The regression parameter is the transition effect and is expressed in logits or as an odds ratio. The effect of the transition on the distal outcome is estimated with the LTA model.

The steps taken to compile these results and summarize the findings is discussed in the Summarizing Conditions section (below).

Empirical Conditions

In order for the current methodological study to reflect conditions in applied research, a random selection of empirical studies using LTA was evaluated. A search using the keywords latent transition analysis was conducted using the PsycINFO database, with the year of publication being limited to 2006-2016. Only peer-reviewed, full-text articles were included in the search. It should be noted that searching a large database does not guarantee that all relevant research will be accounted for, such as unpublished works or theses and dissertations.

The initial search returned 139 articles. A list of ID numbers (1 through 139) was created in a Microsoft Excel spreadsheet and assigned to the returned articles. Next, the Excel RAND function was used to generate a random number for each of the 139 articles. The spreadsheet was then ordered by the random numbers from smallest to largest. The 25 articles atop the newly ordered spreadsheet were reviewed and empirical conditions were noted for each.

The sample sizes for these 25 LTA studies ranged from just over 100 to almost 15,000 participants. Number of indicators ranged from 3 to 17, and the majority of studies included two timepoints. The modal number of classes was 3; however, the true number of underlying latent class structure cannot be verified from these reports. The five-number summaries of empirical sample sizes and number of indicators are presented in Table 3.1.

Table 3.1. Five-number summaries of empirical conditions

Statistic	Sample Size	Indicators
Minimum	109	3
Q1	477	5
Median	658	6
Q3	1600	8.75
Maximum	14718	17

Simulated Conditions

The population structure for the current study is an LTA model with an established number of classes, class prevalences, transition effect, and effect sizes of the distal outcome. Local indepdence holds for all populations. All models will also have a two-class structure. Although the majority of the reviewed applied literature selected either three or four classes, the current study is one of the first methodological examinations of distal effect size estimation in an LTA model. As such, simple transition patterns are desirable.

All models will include five indicators, as the majority of the reviewed studies included either five or six. The indicator response probabilities for the latent classes will not be manipulated in this study as numerous studies have investigated its effects on LCA models. The specification of response probabilities will align with what K. L. Nylund, Asparouhov, and Muthén (2007) refer to as simple structure. Within simple structure, response probabilities for each indicator are either relatively high or relatively low for each class. In line with simple structure in factor analysis, each item is distinctly indicative of one class only. The response probabilities used in this study are presented in Table 3.2.

Table 3.2. Class-specific response probabilities used for five-indicator models

Item	Class 1	Class 2
1	.85	.15
2	.85	.15
3	.85	.15
4	.15	.85
5	.15	.85

The design for this study includes three sample sizes, three class prevalences, three transition effects, and three distal outcome effects. Given that the design is fully crossed, there are 81 total cells (3 sample sizes 3 class prevalences 3 transition effects 3 distal outcome effects = 81) in the design.

Sample Sizes

Based on the review of empirical research, sample sizes for simulated data sets will be 500, 1000, and 1500.

Class Prevalences

Class prevalences will be specified in the population structure. There will be three pairs of class prevalences specified at Time 1: .50/.50, .70/.30, .90/.10. These values were chosen based on their representativeness and coverage of the spectrum of class prevalences of two-class models in applied literature.

Distal Outcome Effects

There will be three distal outcome effect sizes in line with conventional recommendations (Cohen, 1988): small (d = .2), medium (d =.5), and large (d =.8).

Latent Transition Parameters

There are multiple ways to conceptualize and thus specify the transition parameter in LTA. First, one may specify the class prevalences for Times 1 and 2 in addition to specifying the transition parameter. Second, one may simply specify the class prevalences at Time 1 and the transition parameter and allow the class prevalences at Time 2 result from the transition. The advantage of the former option is that many class prevalence and transition effects can be examined. The disadvantage of the former option is that the design balloons very quick in terms of the conditions needed to be simulated (i.e., cells) and some of the combination may not be possible. For example, if three sets of class prevalences were specified for each time point along with, say, three transition parameters, this requires 27 combinations of design factors, some of which may not be estimable or realistic. The advantage and disadvantage for the latter option are the reverse. That is, the advantage is that the design is more manageable, remains squarely focused on the transition parameter, and, most importantly, more accurately approximates the practice of applied researchers. The disadvantage is the fewer transition combinations are examined.

Given the scope and purpose of the current study, we selected the latter option. Given that the latent class variables at Times 1 and 2 are nominal, the transition parameter can be thought of as an odds ratio. We will generate three transition parameters that approximately follow Cohen (1988) and Chen, Cohen, and Chen (2010) recommendations: small (OR = 1.5), medium (OR = 3.5), and large (OR = 7). Using an odds ratio of 1.5 as an example, the interpretation within LTA is that the odds of being in class 1 at Time 2 conditioned on being in class 1 at Time 1 are 1.5 times the odds of being in class 1 in Time 2 conditioned on being in class 2 at Time 1. In other words, it is relatively more likely for those in class 2 at Time 1 to transition to class 1 at Time 2 than for those in class 1 at Time 1 to transition to class 2 at Time 2.

Number of Replications

The number of replications for each cell of the design is 500. This number of replications was selected in order to obtain a stable estimate of the distal outcome effect between samples within the same condition. We expect relatively little sampling error so the choice of 500 replications is sufficient to obtain stable estimates of the effects of primary interest (i.e., transition parameter, distal outcome effect recovery).

Summarizing Conditions

The LTA model used to generate the data will be fit to the simulated data in order to test model parameter recovery. The outcome of interest is recovery of the distal outcome effect size. This will be investigated by measuring raw bias (RB), which is simply the difference between the specified population parameter and that which is estimated in the simulation. First, the estimated values are averaged across each cell, with each cell representing a unique combination of simulation conditions. Then, the difference is calculated between the absolute value of the estimated effect size and the true parameter value. The absolute value of the estimate is used in order to combat label switching that sometimes occurs during simulation of LCA and LTA models. This calculation of RB is written as:

$$|\hat{\theta}| - \theta \tag{3.3}$$

where $\hat{\theta}$ is the parameter estimate and θ is the true parameter value. Positive RB values represent an overestimation of the true parameter, whilst underestimation is revealed in negative RB values. In order to more efficiently evaluate the results of the current study, factorial analysis of variance (ANOVA) will be used to assess main and two-way interactions with raw bias as the dependent variable and all included design factors as independent variables (Bandalos & Leite, 2013). Effect size estimates (η_P^2) will represent the magnitude of the effect of each design factor on the raw bias for the distal outcome estimation. In line with recommendations from previous research

(Bandalos, 2002; Cohen, 1988), effect sizes will be deemed practically significant if they meet or exceed that which is considered to be a large effect size (.14). ANOVA models will be estimated using the aov() command in R.

Summary

The current study is purposed to evaluate model recovery of a distal outcome effect in LTA in order to expand methodological research as LTA becomes an increasingly popular method of analysis in applied settings. While all models will include a two-class structure with two timepoints and five indicators, conditions will include empirically-based variations in sample sizes, Time 1 class prevalences, transition parameters, and distal outcome effect sizes. Fully crossing the conditions will result in 81 cells, with 200 replications being produced for each cell. Raw bias of the distal outcome effect size estimation will be reported for each condition, with positive values representing overestimation and negative ones representing underestimation of the true parameter.

Monte Carlo simulation will be carried out within Mplus, followed by an evaluation of whether the models estimated from generated data sets contain accurate estimates of the specified population structure. Estimation and compiling of output will be automated using the MplusAutomation package within R. Lastly, ANOVA will be applied to these compiled data sets using R to assess the effect each of the design factors has on parameter recovery in LTA models with a distal outcome.

CHAPTER FOUR

Results

The current study is purposed to evaluate how design factors may affect model recovery in an LTA model with a distal outcome. Published, empirical research articles were examined to determine the simulation conditions that would provide results that would be most relevant to applied settings. All models were held to a two-class solution with five indicators, and the varying design factors included sample size (500, 1000, 1500), class prevalence at Time 1 (.50/.50, .70/.30, .90/.10), distal outcome effect size (.2, .5, .8), and the transition parameter (1.5, 3.5, 7). Fully crossing these factors produced 81 cells, each of which was replicated 200 times using Monte Carlo simulation within Mplus (L. K. Muthén & Muthén, 2017). Raw bias (RB) was calculated to measure the discrepancy between the true distal outcome effect size for each condition and those which were estimated from the simulated data sets. ANOVA was then utilized to test for main effects of and interactions between design factors, with RB as the dependent variable. Eta-squared estimates were collected for these effects, and those design factors with effect sizes greater than .14 were considered to have practically significant impacts on model recovery.

Nonconvergence and Overestimation

Of the 16,200 simulated data sets, four models did not converge. All four of these data sets had a sample size of 1000 and the largest transition parameter of 7. Regarding class prevalence at Time 1, three of the four were specified to be split .50/.50 while the true parameter of the fourth was .70/.30. Half of the models had a small true distal effect size (d = 0.2) while the other half was specified as medium (d = 0.5).

Additionally, transition parameters for all models were specified to be 1.5, 3.5, or 7. Upon reviewing the generated parameter estimates, several cells were found to have transition parameter estimates that greatly exceeded the true parameters, with multiple estimates reaching into the tens of thousands (Range: 10.03 − 98,902.56). Among these data sets, all design factors were present in approximately even proportions aside from the transition parameter, which was specified to be the largest parameter in 99.7% of cases. Thirteen additional cases had unspecified transition parameter estimates. Thus, all 2,212 replications with either unspecified estimates or estimates greater than 10 as well as the four models that did not converge were removed from the analyses to avoid inappropriate distortion of results. This removal left 13,989 cells to be analyzed.

Parameter Recovery

As previously stated, parameter recovery of the distal outcome effect size was assessed via RB. ANOVA was used to identify any main effects of and two-way interactions between design factors, with RB as the dependent variable. Analyses revealed that all design factors had partial effect size estimates (η_P^2) less than .14, meaning that parameter recovery was not significantly biased by the included variations of design factors. Table 4.1 shows the ANOVA results, while descriptive statistics can be found in Table 4.2. The mean RB across all cells was 0.0 ($SD = 0.11$).

Sample Size

The RB across the three sample sizes centered approximately on zero, which indicated no bias, on average. The standard deviations were also similar but, as expected, decreased as sample size increased. The empirical 95% confidence intervals for sample sizes of 500, 1000, and 1500 were -0.25 to 0.28, -0.19 to 0.19, and -0.16 to 0.16, respectively. The precision of the estimates increased with larger sample sizes.

Table 4.1. ANOVA with raw bias as the dependent variable

Source	df	SS	MS	F	p	η_P^2
Sample size (SS)	2	0.05	0.03	2.46	.086	3.52e^{-4}
Class prevalence (CP)	2	0.01	0.003	0.30	.743	4.20e^{-5}
Transition parameter (TP)	2	0.02	0.01	0.66	.518	9.40e^{-5}
Effect size (ES)	2	0.11	0.06	5.09	.006*	7.29e^{-4}
SS × CP	4	0.13	0.03	2.84	.023*	8.14e^{-4}
SS × TP	4	0.04	0.01	.087	.481	2.49e^{-4}
SS × ES	4	0.10	0.02	2.16	.071	6.18e^{-4}
CP × TP	4	0.03	0.01	0.57	.687	1.63e^{-4}
CP × ES	4	0.14	0.04	3.22	.012*	9.21e^{-4}
TP × ES	4	0.04	0.01	0.94	.442	2.68e^{-4}
Residual	13956	154.99	0.01			

df: degrees of freedom; SS: sum of squares; MS: mean square; F: F value; p: p value; η_P^2*: partial effect size.*
** denotes statistical significance at the .05 level.*

Class Prevalence

The RB across the three class prevalences centered approximately on zero, which indicated no bias, on average. The standard deviations were also similar but, as expected, increased as the class prevalences became more discrepant. The empirical 95% confidence intervals for class prevalences of .5/.5, .7/.3, and .9/.1 were -0.16 to 0.17, -0.19 to 0.18, and -0.25 to 0.28, respectively. The precision of the estimates decreased as the class prevalences were more discrepant.

Transition Parameter

The RB across the three transition parameters centered approximately on zero, which indicated no bias, on average. The standard deviations were also similar but, as expected, increased as the transition parameter increased. The empirical 95% confidence intervals for transition parameters of 1.5, 3.5, and 7.0 were -0.18 to 0.18, -0.20 to 0.13, and -0.25 to 0.26, respectively. The precision of the estimates was slightly lower when the transition parameter was largest.

Distal Effect Size

The RB across the three distal effect sizes centered approximately on zero, which indicated no bias, on average. The standard deviations were also similar, as expected, across distal effect sizes. The empirical 95% confidence intervals for distal effect sizes of 0.2, 0.5, and 0.8 were -0.18 to 0.23, -0.23 to 0.22, and -0.22 to 0.22, respectively. The precision of the estimates were roughly equal across effect sizes.

Summary

The analyses performed in the current study were intended to reveal any biases in parameter recovery of an LTA model with a distal outcome. All were two-class models with five indicators, and manipulated design factors included sample size, class prevalence at Time 1, the transition parameter, and the distal outcome effect size. Raw bias was used to measure the discrepancy between the true distal outcome effect size and that which was estimated. ANOVA revealed no practically significant effects of the included design factors on model recovery. Further discussion of these results can be found in Chapter Five.

Table 4.2. Raw bias descriptive statistics for each level of each design factor

	Sample size			Class prevalence			Transition parameter			Effect size		
	500	1000	1500	.5/.5	.7/.3	.9/.1	1.5	3.5	7	0.2	0.5	0.8
Mean	0.000	-0.002	-0.002	-0.000	-0.001	-0.000	-0.002	0.000	0.000	0.003	-0.002	-0.003
SD	0.13	0.10	0.08	0.08	0.09	0.13	0.09	0.11	0.13	0.10	0.11	0.11
Min	-0.72	-0.50	-0.39	-0.39	-0.40	-0.72	-0.30	-0.49	-0.72	-0.20	-0.09	-0.72
Max	0.60	0.50	0.39	0.32	0.48	0.60	0.35	0.59	0.60	0.59	0.50	0.60

Values in table depict the raw bias between true and estimated distal outcome effect sizes

SD: standard deviation; Min: minimum value; Max: maximum value

CHAPTER FIVE

Discussion

The purpose of this study was to evaluate the effect of various design factors on parameter recovery of a distal outcome effect within an LTA framework. While methodological research has been performed on LCA with distal outcomes (Asparouhov & Muthén, 2014b; Bakk & Vermunt, 2016; Clark & Muthén, 2009; Lanza et al., 2013) and even LTA with other auxiliary variable such as covariates (Asparouhov & Muthén, 2014a) this study is the first known to address model recovery in LTA models with distal outcomes. Including distal outcomes in classification research is becoming increasingly popular in applied settings, as this addition enables predictions to be made based on class membership. Findings from the current study speak to the ability of LTA models to accurately estimate the distal outcome effect in applied settings.

Parameter Recovery

Four design factors, each with three levels, were specified in this study. These included sample size (500, 1000, 1500), class prevalence at Time 1 (.5/.5, .7/.3, .9/.1), distal outcome effect size (.2, .5, .8), and the transition parameter (1.5, 3.5, 7). Additionally, all models were characterized by two-class solutions at two time points, and all had five indicators. Raw bias (RB) of the distal outcome effect size estimation was measured and subsequently analyzed in ANOVA, with these design factors acting as the independent variables. None of the factors exhibited a practically significant effect on parameter recovery ($\eta_P^2 \leq 0.14$). Therefore, these results support the argument that LTA accurately estimates a distal outcome effect size under the conditions specified in the current study.

Sample Size

Classification methods are known to require large sample sizes in order to be able to detect the underlying latent classes (Cho et al., 2010; Collins & Lanza, 2010; Thompson, 2007). Insufficient sample sizes have also been found to contribute to underestimation of standard errors (Bakk & Vermunt, 2016). Within the current study, the question became whether differences in sample size would alter the ability of an LTA model to correctly estimate a distal outcome effect. As previously mentioned, the included levels of sample size did not show a practically significant effect on parameter recovery. However, the smallest included sample size was 500, meaning that these findings can only be generalized to sample sizes between 500 and 1500. The smallest sample size was set at 500 as this closely resembles the first quartile of the distribution of sample sizes utilized in applied research ($Q1 = 477$), which means that 25% of these applied studies contained sample sizes smaller than 477. However, sample size recommendations for LTA should be followed, and the majority of models are going to require more than 500 subjects in order to detect the latent classes and produce trustworthy results. Therefore, the sample sizes included in this study are sufficient.

Class Prevalence

Another concern within classification is that of starkly uneven class prevalences. If one class is much smaller than the other(s), LCC might not be able to detect it under less-than-desirable conditions (Thompson, 2007). Having a rare class, though, is not unrealistic. In the study by K. L. Nylund et al. (2006) reviewed in Chapter Two, one of the three classes contained only six percent of the total sample at one time point. Until this time, methodological research has not addressed how an uneven split might affect distal outcome estimation in LTA. Results from the current study showed no differences in parameter recovery among the three included class prevalences, the values of which were chosen to span virtually all possible class prevalences for a two-class solution.

Transition Parameter

The transition parameter is that which distinguishes LTA methodology from that of LCA. Therefore, while the other design factors had previously been studied in LCA models with distal outcomes (Asparouhov & Muthén, 2014a; Bakk & Vermunt, 2016; Lanza et al., 2013; Vermunt, 2010), the transition parameter had yet to be studied in any capacity relating to distal outcomes. The current study found no effect of the specified transition parameters on model recovery. The chosen odds-ratios were based on recommendations for small (OR = 1.5), medium (OR = 3.5), and large (OR = 7) transition parameters (Chen et al., 2010; Cohen, 1988). Thus, they are virtually representative of the spectrum of possible transition patterns in applied research. Of course, there can be transition parameters that fall outside of these specifications. Findings from this study can only speak to those LTA models with odds ratios between 1.5 and 7.

As mentioned in Chapter Four, several transition parameter estimates far exceeded the true parameter value. These data sets were not included in analyses as they could have greatly skewed our results. The choice to delete these cases was based on the fact that this study is not evaluating the recovery of the transition parameter. Studying the estimation of the distal outcome effect assumes that the other estimation mechanisms are functioning properly. An evaluation of this overestimation issue will occur separately from the current study.

Distal Outcome Effect Size

RB of the distal outcome effect estimation was unaffected by the specified distal outcome effect sizes. This indicates that regardless of the effect size - within the range of included effect sizes - the effect can be estimated accurately assuming that other conditions are adequate. Previous research supports an inverse relationship between effect size and the number of subjects needed for a study. Our finding speaks only to the relationship between effect size and estimation of the effect, although the analyses also revealed no interaction between sample size and effect size.

Future Research

Since this is the first study of parameter recovery in LTA models with distal outcomes, the number of and variability within the design factors were kept to a minimum. As such, results can only be generalized to those models that fit the specified conditions. Future extensions to the current study could include expanding the number of classes, time points, indicators, and distal outcomes in the LTA model. There could also be the inclusion of other auxiliary variables (e.g. covariates) in addition to the distal outcome(s). Also, as mentioned above, estimation of the transition parameter could be studied to seek out that which caused the overestimation of that parameter in so many data sets.

Recommendations & Conclusion

The current study was purposed to investigate recovery of the distal outcome effect in an LTA model. Design factors of interest included sample size, class prevalences, the transition parameter, and the distal outcome effect size. ANOVA was used to seek out any practically significant effects ($\eta_P^2 \geq 0.14$) of these design factors on the raw bias of the distal outcome estimate. The results revealed no practically significant effects, meaning that the LTA model accurately estimated the distal outcome effect size under all specified conditions. Future research could expand upon this study by including different numbers of classes, indicators, time points, distal outcomes, and other auxiliary variables in addition to the distal outcome(s).

REFERENCES

Asparouhov, T., & Muthén, B. (2014a). Auxiliary variables in mixture modeling: Three-step approaches using mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 329–341.

Asparouhov, T., & Muthén, B. (2014b). Auxiliary variables in mixture modeling: Using the bch method in mplus to estimate a distal outcome model and an arbitrary secondary model. *Mplus Web Notes*, *21*(2).

Bakk, Z., Oberski, D. L., & Vermunt, J. K. (2014). Relating latent class assignments to external variables: standard errors for correct inference. *Political analysis*, 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*(1), 272–311.

Bakk, Z., & Vermunt, J. K. (2016). Robustness of stepwise latent class modeling with continuous distal outcomes. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(1), 20–31.

Bandalos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural equation modeling*, *9*(1), 78–102.

Bandalos, D. L., & Leite, W. (2013). The use of monte carlo studies in structural equation modeling research. In R. O. Hancock G. R. & Mueller (Ed.), *Structural equation modeling: A second course* (2nd Ed. ed., pp. 385–426). Greenwich, CT: Information Age Publishing.

Banfield, J. D., & Raftery, A. E. (1993). Model-based gaussian and non-gaussian clustering. *Biometrics*, *49*(3), 803-821.

Bolck, A., Croon, M., & Hagenaars, J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political Analysis*, 3–27.

Chen, H., Cohen, P., & Chen, S. (2010). How big is a big odds ratio? interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in StatisticsSimulation and Computation®*, *39*(4), 860–864.

Cho, S.-J., Cohen, A. S., Kim, S.-H., & Bottge, B. (2010). Latent transition analysis with a mixture item response theory measurement model. *Applied Psychological Measurement*, *34*(7), 483–504.

Clark, S. L., & Muthén, B. (2009). *Relating latent class analysis results to variables not included in the analysis.*

Clogg, C. C. (1995). Latent class models. In G. Arminger, C. C. Clogg, & M. E. Sobel (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* (p. 311-360). New York: Plenum.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences lawrence earlbaum associates. *Hillsdale, NJ*, 20–26.

Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis.* Hoboken, NJ: John Wiley & Sons, Inc.

Everitt, B. S. (1993). *Cluster analysis* (3rd ed. ed.). New York: John Wiley & Sons, Inc.

Hallquist, M. (2011). *Mplusautomation: Automating Mplus model estimation and interpretation.* Retrieved from http://CRAN.R-project.org/package=MplusAutomation (R version 0.4-2)

Heinen, T. (1996). *Latent class and discrete latent trait models.* Thousand Oaks, CA: Sage.

Hunt, L., & Jorgensen, M. (2003). Mixture model clustering for mixed data with missing information. *Computational Statistics & Data Analysis*, *41*(3), 429–440.

Lanza, S. T., Tan, X., & Bray, B. C. (2013). Latent class analysis with distal outcomes: A flexible model-based approach. *Structural equation modeling: a multidisciplinary journal*, *20*(1), 1–26.

Lazarsfeld, P., & Henry, N. (1968). *Latent structure analysis.* Boston, MA: Hougton Mill.

Lazarsfeld, P. F. (1950). The logical and mathematical foundation of latent structure analysis. In (pp. 362–412). Princeton, NJ: Princeton University Press.

McCutcheon, A. L. (1987). *Latent class analysis* (No. 64). Sage.

McLachlan, G. J., & Basford, K. E. (1988). *Mixture models: Inference and applications to clustering.* New York: M. Dekker.

McLachlan, G. J., & Peel, D. (2000). *Finite mixture models.* New York: John Wiley & Sons, Inc.

McNeil Smith, S., & Fincham, F. (2016). Racial discrimination experiences among black youth: A person-centered approach. *Journal of Black Psychology*, *42*(4), 300–319.

Morgan, G. B., Wells, K. E., Andretta, J. R., & McKay, M. T. (2016). Temporal attitudes profile transition among adolescents: A longitudinal examination using mover- stayer latent transition analysis.

Muthén, B. (2003). Statistical and substantive checking in growth mixture modeling: comment on bauer and curran (2003). *Psychological methods*, *8*(3), 369–77.

Muthén, L. K., & Muthén, B. O. (2010). *Mplus: User's guide* (6th ed. ed.). Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (2017). Mplus. *Statistical analysis with latent variables. Version*, *Version 8*.

Nylund, K., Bellmore, A., Nishina, A., & Graham, S. (2007). Subtypes, severity, and structural stability of peer victimization: what does latent class analysis say? *Child development*, *78*(6), 1706–1722.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, *14*(4), 535–569.

Nylund, K. L., Muthén, B., Nishina, A., Bellmore, A., & Graham, S. (2006). Stability and instability of peer victimization during middle school: Using latent transition analysis with covariates, distal outcomes, and modeling extensions. *Unpublished manuscript*.

Nylund-Gibson, K., Grimm, R., Quirk, M., & Furlong, M. (2014). A latent transition mixture model using the three-step specification. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 439–454.

R Development Core Team. (2017). R: A language and environment for statistical computing (Version 3.4.0), url = http://www.R-project.org ed.) [Computer software manual]. Vienna, Austria.

Rubin, D. B. (1987). Multiple imputation for nonresponse in surveys.

Swanson, H. L., Orosco, M. J., & Lussier, C. M. (2012). Cognition and literacy in english language learners at risk for reading disabilities. *Journal of Educational Psychology*, *104*(2), 302.

Thompson, D. (2007). Latent class analysis in sas: Promise, problems, and programming. In *Sas global forum 2007.* Cary, NC: SAS Institute, Inc.

Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political analysis*, *18*(4), 450–469.

Vermunt, J. K., & Magidson, J. (2002). Latent class cluster analysis. In J. A. Hagenaars & A. L. McCutcheon (Eds.), *Applied latent class analysis* (p. 89-106). Cambridge, MA: Cambridge University Press.