Res High Educ (2014) 55:508–526
DOI 10.1007/s11162-013-9324-5

Latent Class Analysis in Higher Education: An Illustrative Example of Pluralistic Orientation

Nida Denson · Marsha Ing

Received: 30 January 2013 / Published online: 11 January 2014

Abstract Although used frequently in related fields such as K-12 education research, educational psychology, sociology, and social survey research, latent class analysis (LCA) has been infrequently used in higher education. This article provides higher education researchers with a pedagogical application of LCA to classify entering freshmen based on their pluralistic orientation. This study utilized data on entering freshmen at a racially diverse institution on the West coast. LCA was used to estimate latent profile probabilities, classify freshmen into latent classes, and relate latent class probabilities to covariates. The findings indicated that a four-class model was the best fitting model: high pluralistic orientation; high-disposition, low-skill; low-disposition, high-skill; and low pluralistic orientation. Similar to previous research, the findings indicated that the probability of being classified into one group versus the other was dependent upon a student's race/ethnicity and intended major. This approach can aid college administrators in their program planning and targeted interventions around issues of diversity.

Keywords Latent class analysis · Latent variable modeling · Pluralistic orientation · Diversity

Introduction

Although used frequently in related fields, latent class analysis (LCA) has only been recently applied in higher education (e.g., Pastor et al. 2007; Weerts et al. 2013). The purpose of this article is to provide higher education researchers with a pedagogical

N. Denson (✉)
School of Social Sciences and Psychology, University of Western Sydney, Locked Bag 1797,
Penrith NSW 2751, Australia
e-mail: n.denson@uws.edu.au

M. Ing
University of California, Riverside, USA

application of LCA to encourage more widespread use in the field. In this article, we begin by providing a brief overview of the LCA approach, and some applications of LCA in K-12 and higher education. We then provide an illustrative example of LCA to classify students into latent groups based on their pluralistic orientation at the start of college, and examine whether the latent classes were related to students' demographic and background characteristics. Finally, we provide suggestions on how LCA can aid college administrators in their program planning and targeting of interventions.

Overview of the Latent Class Analysis Approach

Latent class analysis, a latent variable modeling approach, is used to classify groups of respondents who are similar on some unobserved construct based on their observed response patterns (Bartholomew 1987; Collins and Lanza 2010; Goodman 1974; Heinen 1996; Muthén 1992, 2001). LCA is sometimes compared to other statistical techniques such as cluster analysis or factor analysis. However, LCA differs from traditional cluster analysis in that LCA is a model-based clustering approach, meaning that the results are not sample dependent and can be replicable in other samples. Unlike in other types of clustering analysis like discriminant analysis (Hand 1981; Huberty 1984; Klecka 1980; Lachenbruch 1975) which has been applied in higher education research (Krosteng 1992; Riggs et al. 1986), in LCA, group membership is not known or observed in the data. Instead, group membership is assumed to be unobserved (latent). Coughlin (2005), for example, notes that discriminant analysis is similar to linear ordinary least squares regression in that it "determines the linear combination of predictor variables that account for the most variation in the group membership outcome" and examines overall model fit in terms of the variance unaccounted for in the model, coefficient statistics, and classification of observations into the observed groups. In Coughlin's (2005) example, retention (retained/not retained) is the observed outcome that is predicted by several observed variables, such as fall semester grade point average and gender. Following this example, in LCA, the outcome would not be observed (or known beforehand) and there would be different ways to evaluate the probability of subjects belonging to particular outcome groups.

Factor analysis and LCA are similar techniques in that both identify unobserved constructs underlying observed responses. In LCA the underlying construct is a categorical variable and in factor analysis it is a continuous variable. Variable-centered approaches (such as factor analysis) identify associations among variables, whereas person-centered approaches (such as LCA) identify groups of individuals who share common attributes (Laursen and Hoff 2006). In variable-centered approaches, it is assumed that the associations among variables are consistent or homogeneous across the population, and these approaches are most appropriate when the aim is to explain variance in outcome variables by a set of predictor variables. In person-centered approaches, it is assumed that the process whereby predictors affect outcome variables is heterogeneous across the population, and these approaches are most appropriate when the aim is to explore individual or group differences in the pattern of responses (for a more detailed discussion comparing the two approaches, see Laursen and Hoff 2006). Masyn (2013) describes latent class analysis as a "person-centered or person-oriented approach (in contrast to variable-centered or variable-oriented)" and cautions of a "practical rejection of the dichotomy between person- and variable-centered approaches" (p. 552–553). Masyn (2013) points out that researchers might "use a person-centered analysis to identify latent classes or groups of individuals characterized by different response patterns on a subset of variables and then use a variable-centered analysis to examine predictors of outcomes (antecedent and consequent correlates) of class membership" (p. 533, e.g. Marsh et al. 2009). Thus, both person-centered and variable-centered approaches serve different functions and may be used within the same study.

Applications of Latent Class Analysis in Education

LCA is used frequently in related areas such as K-12 education, medicine, and sociology (see for example, Harlow et al. 2013; Ho et al. 2012; Ing and Nylund-Gibson 2013; Pastor et al. 2007; Quirk et al. 2013; Weerts et al. 2013). In K-12 education research, for example, Bowers and Sprott (2012) analyzed the Educational Longitudinal Study using LCA to examine a typology of high school dropouts (i.e., *Jaded*, *Quiet*, and *Involved*). They then examined the survey responses four years later to verify their three typology model, and to gain a richer understanding of each dropout group. LCA provided valuable insight into the various types of students at risk for dropping out, and how to better construct future intervention strategies targeted at the different dropout typologies for maximum effectiveness.

In higher education research, Weerts and colleagues (Weerts et al. 2013) analyzed a national dataset from the American College Testing (ACT) program using LCA to classify college students based on their dispositions to civic engagement. The participants were initially assessed as college graduates, and then again one year later. Their LCA analyses revealed four classes of civically engaged students (i.e., *Super Engagers*, *Social-Cultural Engagers*, *Apolitical Engagers*, and *Non-Engagers*). Their study contributed to the literature by providing a more nuanced understanding of how college graduates may be differentiated based on their patterns of civic engagement behavior (as compared to prior studies which have focused mainly on preferences or motives for civic participation) in college. Their analyses also revealed a new type of civically engaged college graduate (i.e., *Social-Cultural Engagers*) not previously identified in the literature.

In another example, Ing and Nylund-Gibson (2013) analyzed data from the Longitudinal Study of American Youth using LCA to examine student attitudes towards mathematics and science. The participants were initially assessed as seventh-graders, and were then followed up 20 years later. The results of their LCA analyses revealed four classes of students (i.e., *Positive*, *Qualified Positive*, *Indifferent*, and *Dim*), and then compared the groups on their demographic characteristics, eighth grade mathematics and science achievement (proximal outcome), and career attainment in science, technology, or mathematics (distal outcome). They suggest future longitudinal research to provide more detailed information about when attitudes start to decline. Such studies can pinpoint exactly when particular groups of students lose interest in science and mathematics, and what factors might contribute to these changes which allows for more targeted interventions. These three examples illustrate a few ways in which LCA has been applied in education settings to identify groups of students, and the implications of identifying those classes for future intervention strategies.

An Illustrative Latent Class Analysis Example in Higher Education

The present study is an illustrative example using LCA to classify student into latent groups based on their pluralistic orientation at the start of college, and to examine whether the latent classes relate to students' demographic and background characteristics. One of the key goals of higher education is to prepare students to work effectively and productively in an increasingly diverse global community (Jayakumar 2008). Pluralistic orientation, an indicator of ones' ability to work effectively with others of diverse backgrounds, being open to new ideas and different perspectives, and being empathetic with other perspectives (e.g., AAC and U, Association of American Colleges and Universities 1995; Hurtado et al. 2002) is an important attribute of our future world citizens, and has been a topic of increased interest in higher education (Engberg and Hurtado 2011; Hurtado et al. 2008; Hurtado and Ponjuan 2005).

Brief Literature Review on Pluralistic orientation

Recent research has shown that there are some group differences in students' pluralistic orientation. Some groups of students have more positive perceptions of the benefits of working with others with different perspectives compared to other groups of students, but other groups do not differ in their perceptions. For example, some research has shown that there are no gender or socioeconomic status differences in students' pluralistic orientation (Engberg et al. 2003; Hurtado and Ponjuan 2005), but that there are differences across students from various racial/ethnic groups and disciplinary contexts (Engberg 2007; Engberg and Hurtado 2011).

Engberg and Hurtado (2011) examined racial/ethnic group differences in students' pluralistic orientation over the first two years of college utilizing data from the *Preparing Students for a Diverse Democracy* project, a longitudinal study of undergraduate students at 10 public universities. Their study found that at the start of college, there were no differences in pluralistic orientation among the racial/ethnic groups (i.e., Whites, Asians, Latinos/as, and Blacks). However, by the end of their second year, there were significant differences between the groups in terms of their pluralistic orientation. In particular, Asian students reported lower scores on pluralistic orientation as compared to the three other groups of students, while the Black, Latino, and White students showed positive increases in their pluralistic orientation over the course of the first two years of college. Thus, their findings illustrate how a student's pluralistic orientation can change during the undergraduate years and emphasize the importance of examining racial/ethnic group differences.

Another study by Engberg (2007) examined the influence of undergraduate student experiences on students' pluralistic orientation across disciplinary academic majors. His study showed that while positive interactions with diverse others had a significant, positive relationship with students' pluralistic orientation at the end of the second year of college, this relationship was strongest for students in the life sciences and business disciplines. In addition, the racial diversity of the student body proved to have an indirect influence on students' second year pluralistic orientation via positive interactions with diverse others. And, these relationships were most pronounced for students in the arts/humanities and engineering disciplines. This study's findings make clear that the institutional context in combination with experiences with diversity can have a positive impact on students' future pluralistic orientation, and that these effects can vary by disciplinary context as well. Taken together, it is becoming apparent that identifying and developing undergraduate students' pluralistic orientation is a priority in the twenty-first century.

Methods

Sample

For this study, we utilized precollege assessments from a national survey of entering freshmen at one racially diverse institution on the West coast. This institution is designated as a Hispanic-Serving Institution, defined by the United States Department of Education as an institution with an enrollment of undergraduate full-time equivalent students that is at least 25 % Hispanic at the end of the award year immediately preceding the date of application. This study utilized freshmen data from Fall 2009, which included 32 % Chicano/Latino, 41 % Asian, 8 % African-American, 15 % White/Caucasian, with the remaining 4 % being Native American, unknown ethnicity, or international student.

Survey

We analyzed student responses from the University of California, Los Angeles (UCLA) Cooperative Institutional Research Program (CIRP) Freshman Survey. The survey was designed in 1966 by Alexander Astin in collaboration with the American Council on Education to measure "established behaviors in high school; academic preparedness; admission decisions; expectations of college; interactions with peers and faculty; student values and goals; student demographic characteristics; and concerns about financing college" (see Dey et al. 1991). The annual survey data collection is organized by the UCLA Higher Education Research Institute (HERI, Sax et al. 2004). Surveys were administered by the institution during orientation for first-year students before the start of fall semester. The high response rate (99 %) was attributed to high attendance at orientation and proctoring of survey administration by representatives from the student affairs department. Participation in the survey was voluntary and anonymous. Participants did not receive any financial compensation for completed surveys and no identifying information was attached to their survey responses. After the surveys were completed, hard copies were sent to the HERI processing center. After data analysis, HERI provided the institution with a de-identified dataset of their student responses, along with a profile report of the institution's overall findings as compared to similar institutions.

From the survey, we selected five items related to the concept of "pluralistic orientation" which is defined as the "skills and dispositions appropriate for living and working in a diverse society" (Engberg 2007; Engberg and Hurtado 2011; Sharkness et al. 2010): ability to see the world from someone else's perspective, tolerance of others with different beliefs, openness to having my own views challenged, ability to discuss and negotiate controversial issues, and ability to work cooperatively with diverse people. The ability to see the world from someone else's perspective and tolerance of others with different beliefs can be considered dispositions, as they are cognitive in nature. The ability to discuss and negotiate controversial issues and work cooperatively with diverse people can be considered skills, as they are actions or behaviors. Openness to having my own views challenged is somewhat of a hybrid between a disposition and a skill. For some, being open to challenging views may be a passive and reflective process, while for others, being open to challenging views may mean engaging and interacting with the person who is challenging your views. For each of these items, students were asked to rate themselves as compared with the average person their age: lowest 10 %, below average, average, above average, highest 10 %). We dichotomized student responses into the "highest 10 %"

versus all other responses (lowest 10 %, below average, average, above average) due to the low percentages of students endorsing the other response categories.[1]

We included the following four covariates based on prior research which has shown group differences in views and attitudes about diversity more broadly (Engberg 2007; Engberg and Hurtado 2011; Sax and Arredondo 1999): gender, race/ethnicity, socioeconomic status, and intended major. Gender was a dichotomous variable (0 = male; 1 = female). Race/ethnicity was also coded as dichotomous variables to represent students who were African American, Asian, Hispanic, Multiracial, Other, and White/Caucasian (reference group). We operationalized socioeconomic status in terms of student self-reports of their parents' income level. We collapsed these categories into three categories (following Chang, Park, Lin, Poon, and Nakanishi, Chang et al. 2007) where low income is <\$60,000, middle income is \$60,000–149,999, and high income is >\$150,000 (reference group). Following the categorization of majors by Engberg (2007), the academic major groupings of a list of 85 different undergraduate major fields were collapsed into six different general categories: (1) arts and humanities (reference group); (2) life sciences (biological/agricultural sciences and health professions); (3) business; (4) social sciences (including cultural/ethnic studies); (5) engineering (including math/physical sciences and computer sciences); and (6) education and social work.

Latent Class Analysis Approach

The latent class analysis approach includes identifying variables that indicate the classes well, estimating the model for different numbers of classes, and then relating class probabilities to covariates and distal outcomes. If the data is available, it is recommended to conduct a split-sample cross-validation or to replicate the model using different data as part of the approach to validate the results of the LCA (Collins et al. 1994; Collins and Lanza 2010; Cudeck and Browne 1983; Masyn 2013; Pastor et al. 2007). The steps in the split-sample cross-validation approach include: (1) randomly splitting the sample into two roughly equally sized subsamples (calibration sample, validation sample); (2) carrying out the latent class analysis for the first subsample (calibration sample); (3) retaining the model parameter estimates for the final model from the first subsample; (4) fitting the final model to the second subsample (validation sample) by fixing all parameters to the estimated values from the final model from the first subsample; (5) evaluating the overall fit of the model for the second subsample with the fixed parameters; (6) fitting the final model to the second subsample allowing all parameters to be freely estimated; and (7) using a nested-model likelihood ratio test to compare the fit of the second subsample using the fixed parameters with the freely estimated parameters. The model fit for the second subsample is

[1] For example, only 4 students (<1 %) indicated that they were in the lowest 10 % in terms of their ability to see the world from someone else's perspective. Due to the low number of respondents endorsing particular categories and an interest in identifying students who were the most positive in their opinions about their abilities, we decided to dichotomize the response options. Continuous response options are also possible using a latent profile analysis (LPA). LPA is an extension of LCA (see Masyn 2013 for additional details). Both are used to classify respondents into latent groups based on their observed response patterns. The difference between LPA and LCA is that LCA uses dichotomous items and thus models the probability of endorsing an item, while LPA uses continuous measures and thus models the means and variance of items to classify respondents into categorical latent groups. We ran the analysis both ways, using the original continuous response options and compared them to the re-coded dichotomous response options. The results were identical in terms of the number of classes identified. For pedagogical purposes, we focus on LCA; however, interested readers can refer to Pastor et al. (2007) for an excellent example of LPA in higher education.

examined and then the fit with the fixed parameters is compared against the freely estimated parameters (Masyn 2013), p. 572. If there is no significant decrease in fit, the model validates the final solution and is considered "stable" across the subsamples. In this study, we followed these procedures to conduct a split sample validation.

Identify Variables

The first step is to identify items that characterize the classes. Prior research and confirmatory factor analysis can help to identify which items are good measures of the pluralistic orientation construct. The items that are determined to be the best indicators are then included in the latent class analysis.

Estimate Models

The next step is to estimate the LCA model for a different number of classes. For each model, respondents are assigned a probability of being in a particular group, given his or her response pattern. Students have a partial probability of being in each of the identified groups. Students may have a higher probability of being in one class as compared to another class. Muthén likens the process of assigning probabilities to factor scores in factor analysis with the following equation with r items and K classes:

$$P(c = k|u_1, u_2, \ldots, u_r) = \frac{P(c = k)P(u_1|c = k)Pu_2|c = k)\ldots P(u_r|c = k)}{P(u_1, u_2, \ldots, u_r)} \tag{1}$$

where P = probability of being in class k give a particular response pattern; c = class, where $c = 1, \ldots, K$; k = class; u_{r} = item number.

The probability of being in a particular class is based on the joint probability of responses to the items.

To evaluate the LCA models, Masyn (2013) and Nylund-Gibson (2012) recommend using several different statistical criteria. In this study, relative fit of the models was compared using the following criteria: Bayesian Information Criterion (BIC), Adjusted BIC (ABIC), Lo-Mendell-Rubin likelihood ratio test (LMR p value), and the bootstrap likelihood ratio test (BLRT p value; Asparouhov and Muthén 2012; Nylund et al. 2007; Lo et al. 2001; Schwartz 1978; Vuong 1989). Lower BIC or ABIC values indicate a better model in comparison with another model with higher BIC or ABIC values. The LMR p value is an indicator of whether the model fits better than the previous model with one less class. A significant p value suggests that the model fits better than the previous model. The BLRT uses bootstrap samples to estimate the distribution of the log likelihood difference and results in p values that are used similarly to those of the LMR. Masyn (2013) describes the Bayes Factor (BF) and the correct model probability (cmP) criteria to evaluate fit. The Bayes Factor asks if given two models, which of the two is more likely to be true (assuming that only one of the two models is true). This value is evaluated for different models with a different number of classes. When the value is higher than 10, there is evidence that this model is the true model. The correct model probability is another way to approximate the correct model probability relative to the different models with different numbers of classes. The model with the largest value is considered the true or correct model.

To interpret the LCA models, Masyn (2013) recommends examining class homogeneity and class separation. Class homogeneity is an indication of whether the items "epitomize each class" (Masyn 2013 , p. 559). High class homogeneity indicates that for a given item, respondents in that particular class are likely to respond similarly to the item. For example, if there is a class (k) with an estimated item (m) probability $(\hat{\omega}_{m|k})$ of 0.90, 90 % of the respondents in that class are likely to endorse that item and 10 % of the respondents in that class are not likely to endorse that item. This item would "epitomize" the class because there is homogeneity in the responses to this particular item. If there was an item with an estimated item probability of 0.50, 50 % of the respondents in the class would endorse the item and 50 % would not endorse the item. This item would not "epitomize" the class because there is not much similarity within the class in terms of endorsing that particular item. Items with estimated item probability >0.70 (high probability of endorsing) or <0.30 (low probability of endorsing) are indications of high class homogeneity (Masyn 2013; Wang and Wang 2012). If there is less consistency in the responses to an item for a particular class (between 0.31 and 0.69), this is an indication that the item is not homogenous for that particular class and item responses are not representative of that particular class.

Class separation is the extent to which the classes are distinguishable or separate from each other. Masyn (2013) describes situations in which it is possible to have high class homogeneity but low class separation; and high class homogeneity and high class separation. In the first hypothetical example, an item might have a 0.95 probability of endorsement for one class and 0.94 for another class. The class homogeneity might be high for this item for both of these classes, however, item endorsement is similar between these two classes and there is not much distinction between these two classes in terms of endorsing this particular item. In the second hypothetical example, it is possible that there is high class separation for two classes with one estimated item probability equal to 0.95 and the other class estimated item probability equal to 0.05. In this situation, there is high class separation because the first class has a high probability of endorsing the item and the second class has a low probability of endorsing that item. To calculate the degree of class separation, Masyn (2013) recommends computing an estimated item endorsement odds ratio using the following formula:

$$\text{Odds ratio}_{m|jk} = \frac{(\hat{\omega}_{m|j}/1 - \hat{\omega}_{m|j})}{\hat{\omega}_{m|k}/1 - \hat{\omega}_{m|k}} \qquad (2)$$

where $\hat{\omega}$ is the estimated item probability for class j and k for item m. This odds ratio is calculated for each item and for each different pairs of classes. For example, if there are three classes, one would compute the odds ratios for class 1 compared to class 2, class 1 compared to class 3, and class 2 compared to class 3. An odds ratio greater than 5 or less than 0.20 is considered an indication of high class separation (Masyn 2013).

Relationship to Covariates

In addition to statistical indicators, Masyn (2013) and Nylund-Gibson (2012) suggest evaluating the model usefulness based on substantively meaningful and substantively distinct classes. There should be both face and content validity of the classes in that they carry some meaning or provide a useful way to characterize the classes. In this study, the relationship between four covariates (gender, ethnicity, socioeconomic status, and intended major) and latent class membership was explored using a series of multinomial logistic regression analyses (Masyn 2013). Class membership was the dependent variable and gender, race/ethnicity, socioeconomic status, and intended major were the independent variables. The multinomial logistic regression analyses provide information on whether prior research and theory about differences in students' pluralistic orientation were predicted by these covariates.

Prior literature suggests that there is no relationship between gender and pluralistic orientation or between socioeconomic status and pluralistic orientation (Engberg et al. 2003; Hurtado and Ponjuan 2005). In other words, females are no more likely than males to have higher or lower pluralistic orientation. Thus, when relating gender to class membership we do not expect to see any gender differences. Prior literature also suggests racial/ethnic group differences in pluralistic orientation (Engberg and Hurtado 2011) and discipline or major (Engberg 2007), so we expect differences in terms of ethnicity and major or discipline.

The models reported in this study were run with Mplus version 7.11(Muthén and Muthén 1998–2012). The Mplus syntax for estimating the final solution is included in the Appendix. More detailed information about how to run these models can be found at the Mplus website and manual.

Results

We randomly split the sample into two equal sizes ($n = 2{,}135$ each sample; Table 1) to carry out the cross-validation steps outlined by Masyn (2013). There were no statistically significant differences between the two samples in terms of endorsement of each of the five pluralistic orientation items (Table 1) or demographic characteristics (for example, there are 46 % male and 54 % female in both the calibration and validation sample). We will refer to the findings of the split-sample cross-validation in the sections below.

Identify Variables

The five pluralistic orientation survey items were selected based on prior research (see Sharkness et al. 2010). Given that the pluralistic orientation factor consisting of the five survey items had already undergone confirmatory factor analysis by UCLA's HERI (Sharkness et al. 2010), it was unnecessary to repeat the procedure here.

Estimate Models

Table 2 presents a summary of the LCA fit indices for one to five latent classes. The BIC and ABIC were the lowest for the four-class solution in both samples; the Bayes Factor (BF) was greater than 10, and the correct model probability (cmP) was the highest in the four-class solution for both samples. When the model parameters from the calibration sample were used on the validation sample, there was good fit to the data, $\chi^2(31, N = 2{,}135) = 41.22$, $p = 0.10$). There was no significant difference in fit between the validation sample with the constrained values and the validation sample with the freely estimated values (log likelihood ratio test $= 26.53$, $p = 0.28$) which indicates that the parameter estimates for the four-class model were not significantly different for the validation and calibration sample. After considering these model comparison statistics and cross-validation results, a four-class model was selected and the latent classes were labeled: high pluralistic orientation; high-disposition, low-skill; low-disposition, high-skill; and low pluralistic orientation. Results from the calibration sample were similar to the validation sample, so to conserve space, only the calibration sample results are presented below.

Table 1 Descriptive statistics

Item[a]	Calibration ($n = 2{,}135$)	Validation ($n = 2{,}135$)	t statistic
1. Compared to the average person your age, rate your ability to see the world from someone else's perspective	21	21	0.16
2. Compared to the average person your age, rate your tolerance of others with different beliefs	34	34	0.00
3. Compared to the average person your age, rate your openness to having my own views challenged	21	21	−0.25
4. Compared to the average person your age, rate your ability to discuss and negotiate controversial issues	24	23	0.49
5. Compared to the average person your age, rate your ability to work cooperatively with diverse people	36	38	0.45

[a] Percent of students who rated themselves as the "highest 10 %"

Table 2 Summary of latent class analysis fit indices with 1–5 latent classes

Sample	Number of classes	Loglikelihood	Number of parameters	BIC	ABIC	VLMR p value	BLRT p value	BF	cmP
Calibration	1	−6,136.74	5	12,311.82	12,295.93	–	–	0.00	–
	2	−4,960.03	11	10,004.40	9,969.45	0.00	0.00	0.00	0.00
	3	−4,882.19	17	9,894.72	9,798.39	0.00	0.00	0.00	0.00
	4	−4,845.25	23	9,866.82	9,736.50	0.00	0.00	520,864,245.24	1.00
	5	−4,842.32	29	9,906.97	9,742.64	0.06	0.00	–	0.00
Validation	1	−6,151.45	5	12,341.24	12,325.35	–	–	0.00	–
	2	−5,015.49	11	10,115.30	10,080.36	0.00	0.00	0.00	0.00
	3	−4,931.06	17	9,992.44	9,938.43	0.00	0.00	0.00	0.00
	4	−4,895.93	23	9,968.18	9,895.11	0.00	0.00	135,637,982.34	1.00
	5	−4,891.66	29	10,005.64	9,913.50	0.07	0.33	–	0.00

Class homogeneity for the four-class solution was examined by the estimated item probabilities (Table 3). The high pluralistic orientation and low pluralistic orientation classes have high homogeneity for all five items which indicates that these two classes are described well by these five items. Item 4 (ability to discuss and negotiate controversial issues) was also a good item for the other two classes. However, the remaining four items (ability to see world from someone else's perspective, tolerance of others with different beliefs, openness to having my own views challenged, ability to work cooperatively with diverse people) did not consistently represent a high degree of class homogeneity for the other two classes. Item 1 (ability to see world from someone else's perspective) was a good item for the low-disposition, high-skill class but not for the high-disposition, low-skill class whereas item 2 (tolerance of others with different beliefs) and item 3 (openness to having my own views challenged) were good items for the high-disposition, low-skill class but not for the low-disposition, high-skill class.

When comparing the estimated item probabilities for pairs of classes, the odds ratios were greater than 5 or less than 0.20 for all of the pairs of classes except the high-disposition, low-skill class and the low-disposition, high-skill class (Table 3). This suggests a high class separation for all five items for the high pluralistic orientation and low pluralistic orientation classes but not for the high-disposition, low-skill class and the low-disposition, high-skill class. The high-disposition, low-skill class and the low-disposition, high-skill class were well separated by two of the five items (tolerance of others with different beliefs and ability to discuss and negotiate controversial issues) only.

Figure 1 presents the item probability profiles for the four latent classes. Over half of entering freshmen were categorized into the "low pluralistic orientation" class (Fig. 1). There were roughly similar percentages of freshmen categorized into the "high pluralistic orientation" and "high-disposition, low-skill" classes. Students in the low pluralistic orientation class were less likely to rate themselves as part of the highest 10 % of others who see the world from someone else's perspective, tolerate others with different beliefs, open to having their views challenged, discuss and negotiate controversial issues, and work cooperatively with diverse people. The high pluralistic orientation group were more likely to rate themselves as the highest 10 % of others across all five items.

Relationship to Covariates

We also explored the relationship between the four classes and four covariates: gender, race/ethnicity, socioeconomic status, and intended major using a series of multinomial logistic regression analyses (Table 4). In terms of gender, there were no differences between males and females in being classified into the high pluralistic orientation group or high-disposition, low-skill group (as compared to the low pluralistic orientation group). However, there were gender differences in terms of females being less likely than males to be classified into the low-disposition, high-skill group (as compared to the low pluralistic orientation group).

In terms of race/ethnicity, Asian students were less likely than White students to be classified in the high pluralistic orientation group, high-disposition low-skill, and the low-disposition high-skill group as compared to the low pluralistic orientation group (reference group). Hispanic students were less likely than White students to be classified in the high-disposition low-skill group, and low-disposition high-skill group as compared to the low pluralistic orientation group (reference group). There were no differences among the African American students, multiracial students, or students who identified as "other" as compared to White students in terms of class membership.

Table 3 Class-specific item response probabilities and odds ratios for four-classes for calibration sample ($N = 2{,}135$)

Item	$\hat{\omega}_{m\|k}$				Odds ratio$_{m\|jk}$					
	1. High pluralistic orientation	2. High-disposition, low-skill	3. Low-disposition, high-skill	4. Low pluralistic orientation	1–2	1–3	1–4	2–3	2–4	3–4
1	**0.76**	0.32	**0.24**	**0.03**	**7**	**10**	**102**	1	**15**	**10**
2	**0.99**	**0.77**	0.34	**0.03**	**30**	**192**	**3,201**	**6**	**108**	**17**
3	**0.81**	**0.25**	0.40	**0.02**	**14**	**7**	**223**	1	**16**	**33**
4	**0.89**	**0.01**	**0.99**	**0.02**	**801**	**0**	**396**	**0**	**0**	**4,851**
5	**0.97**	0.58	0.69	**0.11**	**23**	**15**	**262**	1	**11**	**18**

Item probabilities >0.70 or <0.30 are in bold to indicate high degree of class homogeneity. Odds ratios >5 or <0.20 are in bold to indicate high class separation

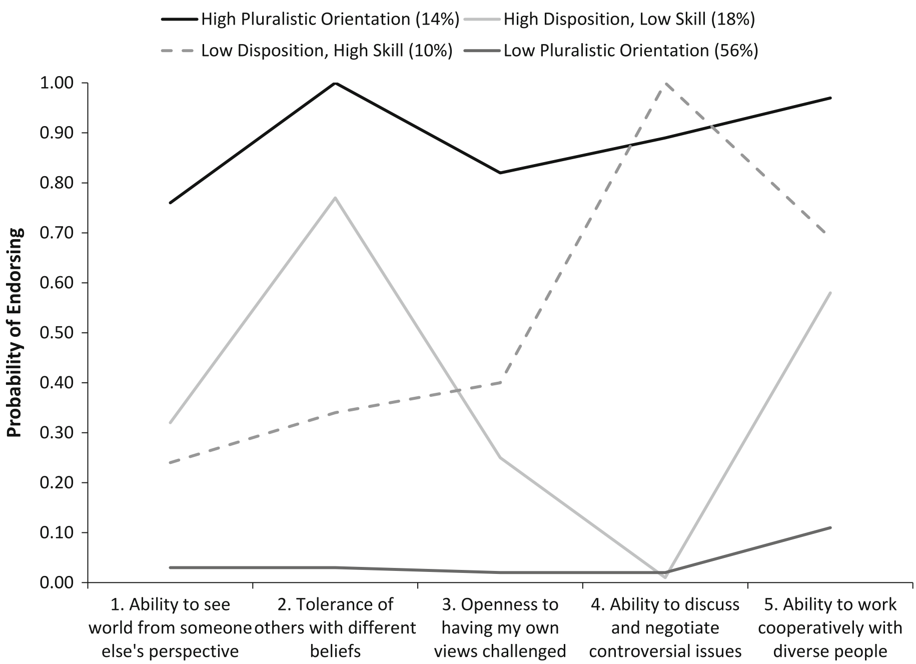

Fig. 1 Item probability profiles for the calibration sample ($N = 2{,}135$)

In terms of socioeconomic status, students from low and middle income backgrounds as compared to students from high income backgrounds were less likely to be in the low-disposition high-skill group as compared to the low pluralistic orientation group (reference group). There were also some differences in predicted class membership by intended major. Students intending to major in the Life Sciences as compared to students in Arts/Humanities were less likely to be in the low-disposition high-skill group as compared to the low pluralistic orientation group (reference group). Students intending to major in Business as compared to students in Arts/Humanities were less likely to be in the high pluralistic orientation group and low-disposition high-skill group as compared to the low pluralistic orientation group (reference group). Students intending to major in Engineering as compared to students in Arts/Humanities were less likely to be in the low-disposition high-skill group as compared to the low pluralistic orientation group (reference group). Students intending to major in Education or Social Work as compared to students in Arts/Humanities were less likely to be in the high-disposition low-skill group as compared to the low pluralistic orientation group (reference group). There were no differences between students intending to major in the Social Sciences as compared to those in Education or Social Work in terms of class membership.

Discussion and Conclusion

This article illustrates an application of latent class analysis in higher education to group individuals into latent classes who are similar on a set of observed variables. The latent class analysis indicated four groups of students: high pluralistic orientation; high-

Table 4 Summary multinomial logistic regression analyses for variables predicting membership into the pluralistic orientation latent classes by gender, race/ethnicity, socioeconomic status, and intended major for calibration sample ($N = 2{,}135$)

	Logit	SE	est/SE	Odds ratio
High pluralistic orientation				
Gender				
Female (vs. male)	−0.11	0.13	−0.89	0.89
Race/ethnicity				
African American (vs. White)	0.22	0.29	0.77	1.25
Asian (vs. White)	−0.77***	0.20	−3.82	0.46
Hispanic (vs. White)	−0.26	0.20	−1.31	0.77
Multiracial (vs. White)	0.09	0.24	0.39	1.09
Other (vs. White)	−0.11	0.32	−0.33	0.90
Socioeconomic status				
Low income (vs. high income)	−0.12	0.21	−0.54	0.89
Middle income (vs. high income)	−0.03	0.22	−0.13	0.97
Intended major				
Life sciences (vs. arts/humanities)	−0.40	0.21	−1.92	0.67
Business (vs. arts/humanities)	−0.68**	0.26	−2.60	0.51
Social sciences (vs. arts/humanities)	0.11	0.24	0.48	1.12
Engineering (vs. arts/humanities)	−0.11	0.23	−0.48	0.89
Education and social work (vs. arts/humanities)	−0.78	0.51	−1.53	0.46
High-disposition, low-skill				
Gender				
Female (vs. male)	−0.02	0.12	−0.15	0.98
Race/ethnicity				
African American (vs. White)	−0.18	0.30	−0.59	0.84
Asian (vs. White)	−0.45*	0.18	−2.52	0.63
Hispanic (vs. White)	−0.39*	0.19	−2.05	0.68
Multiracial (vs. White)	0.11	0.22	0.51	1.12
Other (vs. White)	−0.17	0.31	−0.56	0.84
Socioeconomic status				
Low income (vs. high income)	−0.18	0.20	−0.93	0.83
Middle income (vs. high income)	0.00	0.20	0.05	1.01
Intended major				
Life sciences (vs. arts/humanities)	−0.30	0.19	−1.55	0.74
Business (vs. arts/humanities)	−0.38	0.23	−1.67	0.68
Social sciences (vs. arts/humanities)	−0.13	0.23	−0.55	0.88
Engineering (vs. arts/humanities)	−0.00	0.21	0.00	1.00
Education and social work (vs. arts/humanities)	−1.17*	0.55	−2.10	0.31
Low-disposition, high-skill				
Gender				
Female (vs. male)	−0.50**	0.15	−3.38	−0.60
Race/ethnicity				
African American (vs. White)	−0.08	0.32	−0.25	0.92
Asian (vs. White)	−1.10***	0.22	−5.00	0.33

Table 4 continued

	Logit	SE	est/SE	Odds ratio
Hispanic (vs. White)	−0.70**	0.22	−3.16	0.50
Multiracial (vs. White)	−0.20	0.26	−0.75	0.82
Other (vs. White)	−0.46	0.37	−1.24	0.63
Socioeconomic status				
Low income (vs. high income)	−0.67**	0.21	−3.16	0.51
Middle income (vs. high income)	−0.72**	0.23	−3.15	0.49
Intended major				
Life sciences (vs. arts/humanities)	−0.62**	0.22	−2.79	0.54
Business (vs. arts/humanities)	−0.72**	0.27	−2.64	0.49
Social sciences (vs. arts/humanities)	−0.52	0.28	−1.86	0.59
Engineering (vs. arts/humanities)	−0.61*	0.27	−2.27	0.55
Education and social work (vs. arts/humanities)	−0.54	0.48	−1.05	0.60

Comparison group is ''low pluralistic orientation''

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

disposition, low-skill; low-disposition, high-skill; and low pluralistic orientation. The findings showed that over half of the incoming freshmen at this particular institution were classified into the low pluralistic orientation group, followed by approximately one-fifth into the high-disposition low-skill group, about one-sixth into the high pluralistic orientation group, with the remaining one-tenth into the low-disposition high-skill group. Thus, the majority of incoming freshmen at this one ethnically diverse institution entered college with low levels of pluralistic orientation.

Consistent with previous research, additional analyses indicated that the probability of being classified into one group versus the other was dependent upon a student's race/ethnicity and intended major. There were some differences based on gender and socioeconomic status, but these differences were small and inconsistent. In particular, Asian students entered the university with lower pluralistic orientation as compared to White students. In addition, students intending to major in business as compared to students intending to major in arts/humanities were more likely to be in the low pluralistic orientation class as compared to the high pluralistic and low-disposition high-skill class.

Implications

By identifying students' pluralistic orientation at the start of college, administrators can be better informed about how to target interventions. For example, this information could be used to identify students who are lower in pluralistic orientation (low disposition and low skills) and may benefit from a range of diversity-related activities on campus. This information could also be used to differentiate students who are high in disposition but low in skills from those students who are low in disposition but high in skills as they relate to pluralistic orientation. For example, interventions that target pluralistic orientation *skills* specifically would be most useful for students in the high-disposition, low-skills group. While these students are already predisposed to living in a diverse society such as being open-minded, their skills in this area could be developed further. These students might benefit more from participating in classroom discussions on controversial topics, or by enrolling in a course in which they learn to develop strategies for how to work cooperatively with diverse people. Similarly, interventions that target pluralistic orientation *dispositions* (e.g., ability to see the world from someone else's perspective, tolerance of others with different beliefs) would be beneficial for students in the low-disposition, high-skills group. While these students have the skills to work effectively in a diverse society (e.g., leadership skills), their dispositions could be developed further by participating in perspective-taking activities.

This information could be given to students in the form of feedback on their starting point (baseline) at the beginning of college. It could be used in conjunction with a recommended list of courses that students could choose from if they were interested in developing their pluralistic orientation further. This information could also be used to assist universities in reporting requirements to accrediting commissions for colleges and universities such as the Western Association of Schools and Colleges. At the institution in which we conducted our study, one of the goals of an undergraduate education was to "promote tolerance of the opinions of others and an understanding of the mutual dependence of human beings on each other and on their natural environment." In this way, administrators can use the information gleaned from LCA as part of the assessment and reporting of this general education outcome to the accrediting commission.

Limitations and Future Research

The main limitation of this study is that it was based on self-reported data from a single time point. While it was not possible in this study, the LCA findings could also be supplemented with qualitative methods such as interviews or focus groups to gain a richer understanding of each group. For example, students could be assembled for focus groups about their attitudes toward pluralistic orientation, and perhaps their experiences or environments which influenced their pluralistic orientation. Their responses to the survey could be compared with information gathered from the focus groups. The findings could also be reviewed during meetings with college administrators and faculty to determine the feasibility of the latent classes. Administrators and faculty might have a different perspective of the feasibility of the latent classes based on their interactions with students. Thus, combining statistical and substantive criteria can help validate and interpret the latent classes.

Similar to research conducted by Bowers and Sprott (2012) and Engberg and Hurtado (2011), the current study's findings could also be supplemented by longitudinal data, and how pluralistic orientation is related to outcomes later in life. For example, since it is theorized that a higher pluralistic orientation is an important attribute of our future world citizens, it would be interesting to do a longitudinal follow-up study to examine if and how the students' pluralistic orientation at the start of college was related to their career choices ten years after college. If those students who were in the high pluralistic orientation group chose careers where they interacted frequently with diverse others (e.g., international business), and those students who were in the low pluralistic orientation group chose careers where they were able to avoid interacting with diverse others, the longitudinal findings would provide further support for the various latent classes and would give us a richer understanding of the implications of the latent classes. By providing a brief overview and an illustrative example of latent class analysis in higher education, we hope to encourage more widespread use of this methodological approach in the field.

Appendix

Mplus syntax for latent class analysis predicting four classes.

```
TITLE: Latent Class Analysis.
DATA: file is diversity2009.dat;
VARIABLE: names are subjid divrate1 divrate2 divrate3 divrate4 divrate5;
categorical are divrate1 divrate2 divrate3 divrate4 divrate5;
idvar is subjid;
missing are all (999);
classes = c(4);
ANALYSIS: type = mixture;
starts = 200 50;
process = 4(starts);
OUTPUT: tech11 tech14;
```

References

Asparouhov, T., & Muthén, B. O. (2012). Using Mplus TECH11 and TECH14 to test the number of latent classes (Mplus Web Notes: No. 14). Los Angeles, CA: Mplus.

Association of American Colleges and Universities. (1995). *American pluralism and the college curriculum: Higher education in a diverse democracy*. Washington, DC: Author.

Bartholomew, D. J. (1987). *Latent variable models and factor analysis*. New York: Oxford University Press.

Bowers, A. J., & Sprott, R. (2012). Examining the multiple trajectories associated with dropping out of high school: a growth mixture model analysis. *The Journal of Educational Research, 105*(3), 176–195.

Chang, M. J., Park, J. J., Lin, M. H., Poon, O. A., & Nakanishi, D. T. (2007). *Beyond myths: The growth and diversity of Asian American college freshmen, 1971–2005*. Los Angeles: Higher Education Research Institute, UCLA.

Collins, L. M., Graham, J. W., Long, J. D., & Hansen, W. B. (1994). Cross validation of latent class models of early substance use onset. *Multivariate Behavioral Research, 29*(2), 165–183.

Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. New York: Wiley.

Coughlin, M. (2005). *Applications of intermediate advanced statistics in institutional research*. Tallahassee: Association for Institutional Research.

Cudeck, R., & Browne, M. W. (1983). Cross-validation of covariance structures. *Multivariate Behavioral Research, 18*(2), 147–167.

Dey, E. L., Astin, A. W., & Korn, W. S. (1991). *The American Freshman: Twenty-Five Year Trends*. Los Angeles: Higher Education Research Institute, UCLA.

Engberg, M. E. (2007). Educating the workforce for the 21st century: a cross-disciplinary analysis of the impact of the undergraduate experience on students' development of a pluralistic orientation. *Research in Higher Education, 48*(3), 283–317.

Engberg, M. E., & Hurtado, S. (2011). Developing pluralistic skills and dispositions in college: examining racial/ethnic group differences. *Journal of Higher Education, 82*(4), 416–443.

Engberg, M. E., Meader, E. W., & Hurtado, S. (2003, April). *Developing a pluralistic orientation: A comparison of ethnic minority and White college students*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Goodman, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika, 61*, 215–231.

Hand, D. J. (1981). *Discrimination and classification*. New York: Wiley.

Harlow, D., Nylund-Gibson, K., Taylor, L., & Iveland, A. (2013). Secondary students' views about creativity in the work of engineers and artists: a Latent Class Analysis. *Creative Education, 5*, 315–321.

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

Ho, H. Z., Nylund-Gibson, K., Boutin-Martinez, A., & Tran, C. N. (2012). Profiles of academic risk: Diversity among Asian American students. In R. Endo & X. L. Rong (Eds.), *Educating Asian Americans: Achievement, Schooling, and Identities*. NC: Information Age Publishing, Charlotte.

Huberty, C. J. (1984). Issues in the use and interpretation of discriminant analysis. *Psychological Bulletin, 95*, 156–171.
Hurtado, S., Engberg, M. E., Ponjuan, L., & Landreman, L. (2002). Students' precollege preparation for participation in a diverse democracy. *Research in Higher Education, 43*(2), 163–186.
Hurtado, S., Griffin, K. A., Arellano, L., & Cuellar, M. (2008). Assessing the value of climate assessments: Progress and future directions. *Journal of Diversity in Higher Education, 1*(4), 204–221.
Hurtado, S., & Ponjuan, L. (2005). Latino educational outcomes and the campus climate. *Journal of Hispanic Higher Education, 4*(3), 235–251.
Ing, M., & Nylund-Gibson, K. (2013). Linking early science and mathematics attitudes to long term science, technology, engineering and mathematics career attainment: Latent class analysis with proximal and distal outcomes. *Educational Research and Evaluation, 19*(6), 510–524.
Jayakumar, U. M. (2008). Can higher education meet the needs of an increasingly diverse and global society? Campus diversity and cross-cultural workforce competencies. *Harvard Educational Review, 78*(4), 615–651.
Klecka, W. R. (1980). *Discriminant analysis*. Beverly Hills: Sage.
Krosteng, M. V. (1992). Predicting persistence from the student adaption to college questionnaire: early warning or siren song? *Research in Higher Education, 33*, 99–111.
Lachenbruch, P. A. (1975). *Discriminant analysis*. New York: Hafner.
Laursen, B., & Hoff, E. (2006). Person-centered and variable-centered approaches to longitudinal data. *Merrill-Palmer Quarterly, 52*(3), 377–389.
Lo, Y., Mendell, N., & Rubin, D. (2001). Testing the number of components in a normal mixture. *Biometrika, 88*(3), 767–778.
Marsh, H. W., Ludtke, O., Trautwein, U., & Morin, A. J. S. (2009). Classical latent profile analysis of academic self-concept dimensions: Synergy of person- and variable-centered approaches to theoretical models of self-concept. *Structural Equation Modeling, 16*(2), 191–225.
Masyn, K. (2013). Latent class analysis and finite mixture modeling. In T. Little (Ed.), *The Oxford handbook of quantitative methods in psychology* (Vol. 2, pp. 375–393). Oxford: Oxford University Press.
Muthén, B. O. (1992). Latent variable modeling in epidemiology. *Alcohol Health and Research World, 16*, 286–292.
Muthén, B. O. (2001). Latent variable mixture modeling. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New development and techniques in structural equation modeling* (pp. 1–33). Mahwah: Lawrence Erlbaum Associates.
Muthén, B. O., & Muthén, L. K. (1998–2012). *Mplus User's Guide*. Los Angeles, CA: Muthén and Muthén.
Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in mixture modeling: A Monte Carlo simulation. *Structural Equation Modeling, 14*(4), 535–569.
Nylund-Gibson, K. L. (2012, November) *Latent class analysis in higher education*. Paper presented at the annual meeting of the Association for the Study of Higher Education, Las Vegas, NV.
Pastor, D. A., Barron, K. E., Miller, B. J., & Davis, S. L. (2007). A latent profile analysis of college students' achievement goal orientation. *Contemporary Educational Psychology, 32*, 8–47.
Quirk, M., Nylund-Gibson, K., & Furlong, M. (2013). Exploring patterns of Latino/a children's school readiness at kindergarten entry and their relations with grade 2 achievement. *Early Childhood Research Quarterly, 28*(2), 437–449.
Riggs, M., Downey, R., McIntyre, P., & Hoyt, D. (1986). Using discriminant analysis to predict faculty rank. *Research in Higher Education, 25*, 365–476.
Sax, L. J., & Arredondo, M. (1999). Student attitudes toward affirmative action in college admissions. *Research in Higher Education, 40*(4), 439–459.
Sax, L. J., Hurtado, S., Lindholm, J. A., Astin, A. W., Korn, W. S., & Mahoney, K. M. (2004). *The american freshman: national norms for fall 2004*. Los Angeles: Higher Education Research Institute.
Schwartz, G. (1978). Estimating the dimensions of a model. *The Annals of Statistics, 6*(2), 461–464.
Sharkness, J., DeAngelo, L., & Pryer, J. (2010). CIRP construct technical report. Los Angeles, CA; Higher Education Research Institute. Retrieved December 1, 2012, from http://www.heri.ucla.edu/PDFs/constructs/technicalreport.pdf
Vuong, Q. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica, 57*, 307–333.
Wang, J., & Wang, X. (2012). *Structural equation modeling. Applications using Mplus*. West Sussex: Wiley.
Weerts, D. J., Cabrera, A. F., & Mejias, P. P. (2013). Uncovering categories of civically engaged college students: A latent class analysis. *The Review of Higher Education, 37*, 141–168.