

**Title:** Identify the typology of Early Childhood education through Parents' perception: A Latent class analysis of ECPP:2016

**Abstract:** A child's early years are the foundation for his or her future development, providing a strong base for lifelong learning and learning abilities, including cognitive and social development. Yet, since to date the majority of studies on early childhood focus on children's participation in and barriers to participating in nonparental care arrangements. The purpose of this study is to examine the different typologies of early childhood education from a large nationally representative dataset (ECPP:2016) using latent class analysis (LCA) based on parent's perception. We found five significantly different latent class. Based on this typology, we discuss implications for future early childhood education research and practice.

**Background:** Educators and policymakers are pressed to respond to the current crisis in American public education. Concerns about the crisis in public education have led to the establishment of the National Education Goals (National Education Goals Panel, 1999). The goals represent a strategic plan to enhance learning opportunities for all students by targeting what experts believe to be the most fundamental components of effective intervention. One of the cardinal foci of this strategic national plan is school readiness. This goal and its accompanying objectives highlight the need for quality early childhood educational programs and emphasize the importance of establishing and maintaining substantial parent involvement to promote student learning. This was clearly manifest in 2006 when the California legislature unsuccessfully attempted to pass a law that would make pre-school education mandatory for all children. Meanwhile, according to some studies, children enrolled in preschool programs are more behaved and have higher IQ scores upon enrolling kindergarten and learn quicker than their peers without formal education.

However, there still limit information from the empirical typology analysis about the early childhood education. The conventional typology classifies early childhood education into the classes like child care by relatives (e.g., grandparents), child care by non-relatives (e.g., neighbor), day care center, and preschool program. However, the information and insight we can obtained from this typology is limited since it completely ignores the fundamental differences in parent's perceptions about what factors matter for choosing a child care arrangement. To fill in this gap, we explore a new typology of early childhood education based on the self-reported data from parents.

**Research Question:** There are three research questions that we focus on this study: (1) Does there exist latent classes for the believe of early childhood education from parents? (2) How children and family background related to the estimated latent classes? and (3) Do latent classes effect children's learning?

**Data Sources:** This study is a secondary analysis of Early Childhood Program Participation Survey (ECPP:2016). ECPP:2016 is part of the National Household Education Surveys Program (NHES:2016), which is the flagship household survey program of the National Center for Education Statistics (NCES). NHES:2016 used a nationally representative address-based sample covering the 50 states and the District of Columbia. ECPP:2016 focus on young children's care and education before kindergarten. The surveys typically focus on children's participation in and barriers to participating in nonparental care arrangements; what activities the

family does with the child, such as reading, singing, and arts and crafts; and what the child is learning, such as counting, recognizing the letters of the alphabet, and reading. Parents are the respondents.

The whole survey in this study children in the United States from birth through the age of 5 who were not yet enrolled in kindergarten by 2016. After removing the data set which miss all the value in the indicator variable in LCA model,  $n = 3871$  children are included in this study.

**Methods:** We used the three-step LCA model under the framework of mixture model (Jung & Wickrama, 2008; Nylund, Asparouhov, & Muthén, 2007). MPLUS (Muthén & Muthén, 2007) was used to conduct the latent class analysis (see appendix for MPLUS code). *Figure 1* details the framework of the LCA model. Latent class is first estimated based on 10 indicators about the perception from parents about the main reason for choosing care (i.e., what factors were important to parents when choosing a care arrangement). *Table 1* shows the descriptive statistics of the indicators before transformation. All indicators (originally measured in four level Likert) are dichotomized by setting ‘not at all important’ and ‘a little important’ as 0 and ‘somewhat important’ and ‘very important’ as 1, when input into the LCA model.

In the second step, six covariate variables are used to fit on the estimated latent classes. These covariates are related to the background information of the children and their family (see *Table 4* for the descriptive analysis of the covariates in original scale). Again, we dichotomized all covariates with the threshold as the middle measurement level in the original scale. In particular, the variable about disability is based on 13 specific disability or disturbance items. If the child has at least one condition of disability, he/she will have value of 1. Otherwise, it is transformed as 0 (i.e., do not have any kinds of disability or disturbance).

Finally, to assess the interpretation of each of the estimated latent class as well as to explore the reasons that students have different learning performance at early age. Four learning outcome variables are used as the dependent variable (see *Table 6* for descriptive analysis of the distal outcome in original scale) in the third step with estimated latent class as independent variable. All outcome variables are also dichotomized for interpretation purpose. We need to note that children under the age of 2 (884 observations in total) are not required to answer the questions about children’s learning. Consequently, all missing value because of age will be set as 0, which correspond to the situation that the children do not master these skills.

**Results:** In general, LCA with five latent class have the best model fit. Based on the model fit statistics (see *Table 2*), five-class result has the lowest information criteria values ( $AIC = 30020.140$  and  $BIC = 30215.438$ ) with the highest entropy value of 0.681. The best loglikelihood value has been replicated. However, the evidences from Lo-Mendell-Rubin Adjusted LRT test is not significant, which means we may have a significant improve of model fit if we keep adding more latent classes into the model. But for the purpose of interpretation and generalization, we prefer the parsimonious model result.

Except for the first latent class, probabilities all exceeding 0.8 for the most likely membership in each class indicating separation of individuals across the three latent class model (see *Table 3*). For the first latent class, it has a limitation in distinguish itself with the second and third latent classes. The proportion of five latent classes based on posterior probability are 10.54%, 48.89%, 32.35%, 1.60%, and 6.63% respectively. Some previous empirical studies suggest that each latent class should have a proportion bigger than 10%. However, starting from the model with three latent class, there will always have at least one latent class capture less than 10%. This

partially indicates that a small proportion of sample in this dataset do have a distinct response pattern, which is unlikely to be combine with other common patterns.

In general, ‘reliability’ and ‘learning activity’ are the two most important factors with high agreement. By contrast, ‘rating on the website’ and ‘religious orientation’ have lowest importance with low agreement (see *Table 1*). *Figure 2* details the response patterns of each of the five identified typologies to the 10 dichotomously scaled indicators. The x-axis provides each of the items, while the y-axis presents the proportion of each group that responded “important” or “not important. We named the first latent class as “learning and peer oriented”, parents in this class want to provide their kids with an environment of many kids at the similar age and they can spend a lot of time together to explore new things. They do not pay much attention towards the location, cost, or reliability of the arrangement. Instead, they give relative high importance towards religious orientation. We labeled the parents in the second latent class as “demanding”. They give relative high importance to all indicators compared with the parents from other latent classes. In contrast, the fourth latent class gives relative low importance to almost all indicators compared with other latent classes, except ‘spending time with other kids his/her age’ and ‘recommendation from friend and family’. Thus, we named fourth latent class as “Relaxed”. We named the fifth latent class as “economic & convenience oriented” since parents from this class pay relative higher attention towards the location, cost, and reliability of arrangement. They also wish the arrangement could provide day care. Finally, we named the third latent class as “mixed” since parents from this class have many similar patterns as some other classes. Similar to the “economic & convenience oriented” class, they care about the location, cost and reliability. Meanwhile, they rely more on family and friends’ recommendation than the rating on the website. They do not care about the religious orientation. Different from “economic & convenience oriented” class, they also pay attention towards the learning activity and the opportunities of spending time with other kids, like the “learning and peer oriented” class.

Based on the result from second-step of LCA model, we take “demanding” class as reference group since it capture the most common response pattern (see *Table 5*). Children in “learning and peer oriented” are significantly more likely to come from a high income and white family, which may partially explain why their parents give low importance value on ‘location’, ‘cost’, and ‘reliability’. Children in the “mixed” class are significantly more likely to come from a small and white family. Children in “relaxed” class is significantly more likely to be disable and comes from a big family. This may partially explain why their parents give low importance value on almost all indicators except learning activity and opportunities of spending time with other kids. Finally, children from “economic & convenience oriented” class is significantly more likely to come from a poor and white family. This may partially explain why their parents give high importance value on ‘location’, ‘cost’, and ‘reliability’.

Based on the third-step of LCA model (see *Table 7*), the children from “learning and peer oriented” class have best outcome on all variables, while children from “economic & convenience oriented” class have worst outcome on all variables compared with other latent classes. Comparing the response pattern of these two classes, we could see the importance of ‘learning activity’ and the interaction with ‘other children’ in preschool education. Similarly, children from the “mixed” class have the second-best performance on almost all variable since their parents also pay much more attention to ‘learning activity’ and the interaction with ‘other children’. Another explanation is that children from “learning and peer oriented” class is more likely to comes from the rich family, who are willing to find the good preschool arrangement regardless of ‘cost’, ‘location’, and ‘reliability’. Based on the evidence from Pearson pairwise test (see *Table 8*), children from

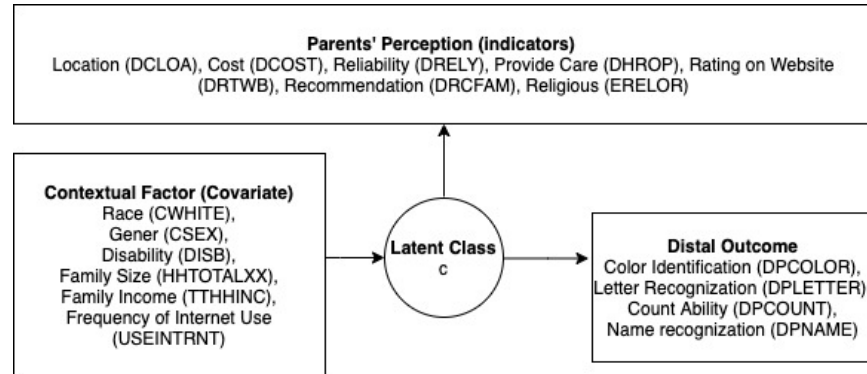
“demanding” class and “relaxed” class have no significant difference on all distal outcomes. Surprisingly, their parents’ perceptions are very different with parents. This partially indicates that it is hard to make really contributing decision when parents give high importance to all factors in reality.

**Conclusion:** These findings are novel and significant extensions of the past early childhood education research for three main reasons. First, our findings give a five distinctive latent class based on parents’ perception about early childhood education. Second, the background information of the kids and family (except gender), to some extent, can explain the difference among the latent classes. Finally, the latent class has significant difference in their learning outcome, which gives some insight to the early childhood education practice. However, this study also has some limitations. For example, the overall model fit in LCA need to be improved and the accuracy of most likely probability in first latent class is low. Meanwhile, the results from the three-step LCA is not correlational rather than casual. The longitudinal data analysis using similar LCA methods can be explore in the future research.

#### Reference

- Jung, T., & Wickrama, K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*, 2(1), 302-317.
- 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: A Multidisciplinary Journal*, 14(4), 535-569.
- Muthén, L. K., & Muthén, B. O. (2007). *Mplus: Statistical analysis with latent variables user's guide* (5th ed.). Los Angeles, CA: Muthén & Muthén.
- Children's Defense Fund. (1997). *Children's Defense Fund Annual Report: Celebrating 25 Years*. Children's Defense Fund, Washington, DC.
- Thomas M., Smith B., Aronstamm Y., Susan P. C., Marianne P., Nabeel A., Mary R. R., & Yupin B., (1996). *The Condition of Education 1996*. U.S. Department of Education Office of Educational Research and Improvement, Washington, DC.
- National Education Goals Pane, (1999). *The National Education Goals Report: Building A Nation of Learners 1999*, Washington, DC.

## Appendix



**Figure 1:** Latent Class Analysis (LCA) model for early childhood education

**Table 1:** Descriptive Analysis of the Indicator variables of LCA

Variables	N	Min	Max	Mean	SD	ECPP:2016
Location of arrangement	3871	1	4	3.434	0.790	CDLOA
Cost of arrangement	3871	1	4	3.247	0.903	DCOST
Reliability of arrangement	3871	1	4	<u>3.813</u>	<u>0.516</u>	DRELY
The learning activity at the arrangement	3871	1	4	<u>3.719</u>	<u>0.563</u>	DLERN
The child spending time with other kids his/her age	3871	1	4	3.550	0.773	DCHIL
The times during the day that this caregiver is able to provide care	3871	1	4	3.560	0.752	DHROP
The number of other children in the child's care group	3871	1	4	3.106	0.893	DNBGRP
Ratings on a website	3871	1	4	<u>2.576</u>	<u>1.135</u>	DRTWEB
Recommendation from friends and family	3871	1	4	3.192	0.967	DRECFAM
Religious orientation of the arrangement	3871	1	4	<u>2.022</u>	<u>1.136</u>	DRELOR

**Table 2:** LCA results and fit statistics for Early Childhood Education

Model	AIC	Adjusted BIC	-Log likelihood	LMR	P-value	Entropy
Two Classes	30756.580	30821.339	15357.290	1665.097	0.000	0.623
Three Classes	30456.748	30555.428	15196.374	318.329	0.009	0.671
Four Classes	30202.608	30335.208	15058.304	273.135	0.000	0.663
<b>Five Classes</b>	<b>30020.140</b>	<b>30215.438</b>	<b>14970.458</b>	<b>173.779</b>	<b>0.000</b>	<b>0.681</b>
Six Classes	30048.916	30220.582	14945.070	50.224	0.000	0.647

**Table 3: Classification Probabilities for the Most Likely Latent Class Membership**

Latent Classes	1	2	3	4	5
1	<b>0.531</b>	0.220	0.236	0.006	0.008
2	0.011	<b>0.843</b>	0.142	0.000	0.004
3	0.033	0.132	<b>0.807</b>	0.002	0.026
4	0.051	0.000	0.076	<b>0.816</b>	0.057
5	0.005	0.030	0.111	0.009	<b>0.845</b>

**Figure 2: Latent class pattern****Table 4: Descriptive Analysis of the Covariates**

Variables	N	Min	Max	Mean	SD	ECPP:2016
Children's Race (white or not)	3871	0	1	0.806	0.396	CWHITE
Children's Gender (male or not)	3871	0	1	0.514	0.499	CSEX
Whether Children Has any Kind of Disability and Disturbance	3871	0	1	0.123	0.327	HDINTDIS, HDSPEECHX, HDDISTRBX, HDDEAFIMX, HDBLINDX, HDORTHOX, HDAUTISMX, HDPDDX, HDADDX, HDLEARNX, HDDELAYX, HDTRBRAIN, HDOTHERX
Number of people living in the family	3871	0	9	3.813	1.188	HHTOTALXX
Frequency of using Internet	3871	0	4	2.719	0.489	USEINTRNT
Total income of the family	3871	0	9	2.550	2.802	TTLHHINC

**Table 5:** Mean and Odd Ratio for Covariates across Latent Class with Latent Class 2 as Reference Group

Covariates	Learning & Peer Oriented (10.54%)			Mixed (32.35%)			Relaxed (1.60%)			Economic & Convenience Oriented (6.82%)		
	mean	Odd Ratio	P-value	mean	Odd Ratio	P-value	mean	Odd Ratio	P-value	mean	Odd Ratio	P-value
Disability	0.131	1.140	0.641	0.184	1.202	0.290	<u>1.081</u>	<u>2.948</u>	<u>0.005**</u>	-0.554	0.574	0.119
Number of People	0.201	1.223	0.292	<u>-0.559</u>	<u>0.572</u>	<u>0.000***</u>	<u>0.879</u>	<u>2.408</u>	<u>0.004**</u>	0.167	1.182	0.335
Internet Use	-0.255	0.775	0.683	0.801	2.229	0.214	-0.060	0.942	0.942	-0.028	0.972	0.969
Family Income	<u>0.426</u>	<u>1.530</u>	<u>0.055</u>	0.053	1.054	0.678	-0.371	0.690	0.241	<u>-1.094</u>	<u>2.985</u>	<u>0.000***</u>
Gender	0.096	1.101	0.598	0.077	1.081	0.499	-0.215	0.806	0.519	0.132	1.141	0.427
Race	<u>1.479</u>	<u>4.390</u>	<u>0.000***</u>	<u>1.008</u>	<u>2.739</u>	<u>0.000***</u>	0.269	1.309	0.428	<u>0.732</u>	<u>2.080</u>	<u>0.002**</u>

**Table 6:** Descriptive Analysis of the Distal Outcome

Variables	N	Min	Max	Mean	SD	ECPP:2016
Color identification	3871	0	2	1.373	0.863	DPCOLOR
Letter reorganization	3871	0	3	1.435	1.214	DPLETTER
Count Ability	3871	0	5	2.274	1.635	DPCOUNT
Whether kid can write the first word even backward	3871	0	1	0.385	0.486	DPNAME

**Table 7:** Mean and Standard Deviation of Outcome across Latent Classes

Outcome Variables	Learning & Peer Oriented (10.54%)		Demanding (48.89%)		Mixed (32.35%)		Relaxed (1.60%)		Economic & Convenience Oriented (6.82%)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Color identification	<u>0.903</u>	0.027	0.776	0.013	0.865	0.016	0.715	0.075	<u>0.654</u>	0.053
Letter reorganization	<u>0.722</u>	0.039	0.588	0.016	0.596	0.022	0.589	0.083	<u>0.385</u>	0.058
Count Ability	<u>0.810</u>	0.036	0.664	0.015	0.674	0.021	0.527	0.084	<u>0.514</u>	0.058
Whether kid can write the first word even backward	<u>0.52</u>	0.054	0.516	0.016	0.514	0.022	0.452	0.084	<u>0.307</u>	0.041

**Table 8:** Pearson Pairwise Test of Mean with Latent Class 2 as Reference Group

Outcome Variables	Learning & Peer Oriented (10.54%)		Mixed (32.35%)		Relaxed (1.60%)		Economic & Convenience Oriented (6.82%)	
	Chi-square	P-value	Chi-square	P-value	Chi-square	P-value	Chi-square	P-value
Color identification	<u>16.195</u>	<u>0.000***</u>	<u>14.470</u>	<u>0.000***</u>	0.645	0.422	4.988	0.026
Letter reorganization	<u>9.084</u>	<u>0.003**</u>	0.074	0.786	0.000	0.998	<u>11.449</u>	<u>0.001***</u>
Count Ability	<u>13.000</u>	<u>0.000***</u>	0.113	0.737	2.556	0.110	6.305	0.012
Whether kid can write the first word even backward	<u>20.625</u>	<u>0.000***</u>	0.005	0.945	0.569	0.451	<u>17.204</u>	<u>0.000***</u>

## MPLUS Code

```
Title: Three-step LCA Model for Early Childhood Education
Data: File = C:\Users\ATS-Workshop\Desktop\data.dat ;
Variable: NAMES = dcloa dcost drely dlern dchil dhrop dnbgrp drtweb
             drecfam drelor rcnow ncnow cpnnowx DISB hhtotalxx
             useintrnt ttlhhinc csex cwhite
             dpcolor dpletter dpcount dpname ;
MISSING = ALL(999) ;
USEVARIABLES = dcloa dcost drely dlern dchil dhrop
              dnbgrp drtweb drecfam drelor ;
CATEGORICAL = dcloa dcost drely dlern dchil dhrop
              dnbgrp drtweb drecfam drelor ;
CLASS = c(5) ;
AUXILIARY = (R3STEP) DISB hhtotalxx
            useintrnt ttlhhinc csex cwhite;
! Command Syntax for step 2
AUXILIARY = dpcolor (BCH) dpletter (BCH)
            dpcount (BCH) dpname (BCH) ;
! Command Syntax for step 3
Analysis: TYPE = MIXTURE ;
          MITERATION=10000;
          STARTS = 25000 250;
          STITERATIONS = 100;
Output: tech10 tech11 tech 14;
Savedata: SAVE = CPROBABILITIES ;
          FORMAT=FREE;
          FILE = C:\Users\ATS-Workshop\Desktop\CPROBS-011.DAT ;
          ESTIMATES = C:\Users\ATS-Workshop\Desktop\MIXEST-011.DAT ;
```