# An Exploratory Latent Class Analysis on Early Coding Attitudes of Fourth Grade Students

By Leiny Garcia • School of Education • University of California, Irvine

## Why Early Coding Attitudes?

- When it comes to CS participation in the US, **Latinx students** are **2x's less likely** to take an AP CS exam than their White and Asian peers when they attend a school that offers it.[1] Latinas compose 18% of all female exam takers despite making up 27% of the female student population.[1]

***We must explore the formation of early CS attitudes among underrepresented youth.***

- **Attitude** is a cognitive mental state associated towards an object that influences behavior.[2] From a socio-cognitive perspective, attitudes is an entity that is influenced by the environment and therefore an indicator of a forming identity.[3]
- **Expectancy-Value Theory (EVT)** is a lens to identify attitude constructs with a focus on achievement-related choices. Positive and negative components of **success expectancies and values (SEV)** on a task directly influence performance, persistence, and task choice.[4]

## Methods

***Sample*** Participants in this study attend a local public school district with a **student population that is predominately Latinx (96.1%) and of low socioeconomic background (87.0%).**[5] District students have been exposed to CS primarily through ad-hoc events like Hour of Code. A total of 12 educators across 6 schools volunteered to administer coding assessments which included a pre-test on coding attitudes. The final sample size is **N=264 of which 137 are female and 127 are male.**

***Measurement*** This study uses the Elementary Student Coding Attitudes Survey (ESCAS) which is a validated measurement for early coding attitudes among pre-adolescent age groups.[6] Using the EVT framework, they identified five latent constructs that pertain to coding:

**Confidence (C) Interest (I) Utility (U)**
**Perception of Coders (P) Social Influences (S)**

***Analyses***

This study conducted a **Latent Class Analysis (LCA)** on the pre-test ESCAS survey responses using the R package poLCA.[7] LCA provides a ***person-oriented approach*** and identify latent classes or subgroups types that differ in response patterns.

Items that were independent of a deep understanding of code was used and converted from a 6 likert scale *(strongly disagree/strongly agree)* to a dichotomous one.

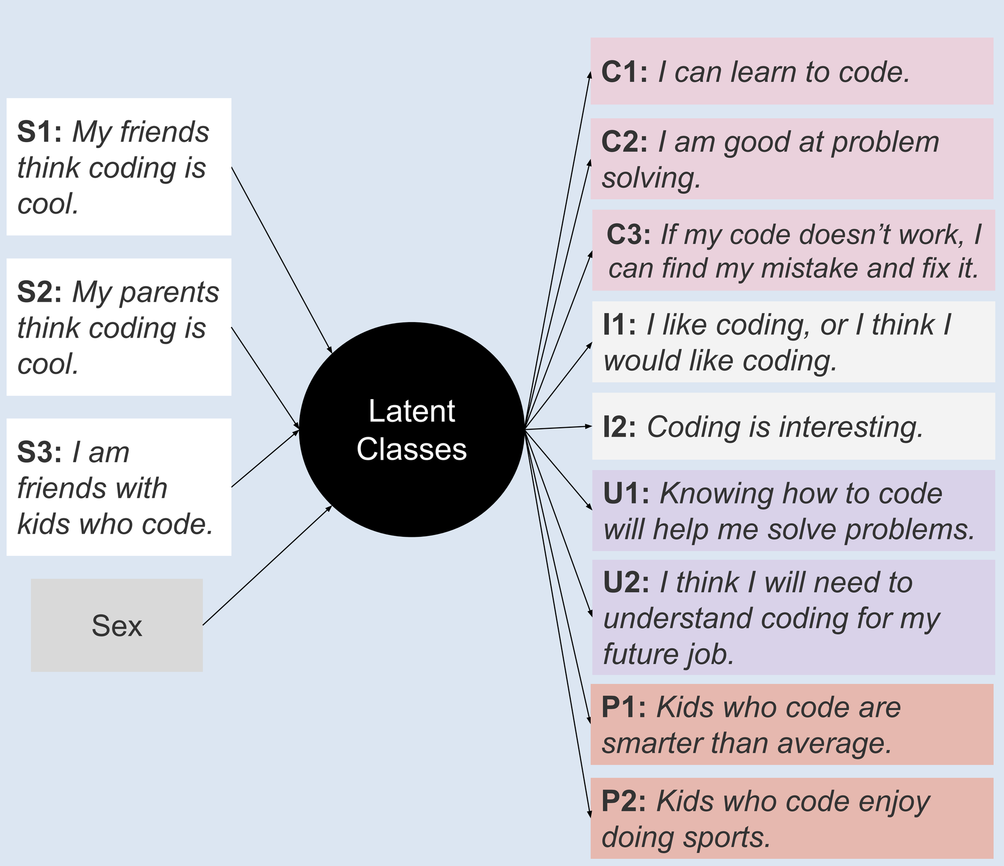



**Figure 1**. Latent Class Model Diagram

**Table 1.** Percentage of Student Responding "Agree" to Coding Attitude Items by Sex

| | C1 | C2 | C3 | I1 | I2 | U1 | U2 | P1 | P2 | S1 | S2 | S3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Total *N=264* | 80.68 | 74.62 | 76.51 | 74.24 | 80.68 | 76.14 | 64.39 | 65.90 | 48.86 | 66.29 | 63.26 | 40.53 |
| Sex | | | | | | | | | | | | |
| Female *N=137* | 79.56 | 70.07 | 78.10 | 75.91 | 81.75 | 74.45 | 62.77 | 60.58 | 48.90 | 63.50 | 68.61 | 40.88 |
| Male *N=127* | 81.89 | 79.53 | 74.80 | 72.44 | 79.53 | 77.95 | 66.14 | 71.65 | 48.82 | 69.29 | 57.48 | 40.16 |

## RQ1: Are there different types of coding attitudes among fourth grade students with little to no experience in CS?

### Class Enumeration

This exploratory LCA study underwent a class enumeration process to determine the best fit model with the following criteria:

- Smallest values for Bayesian Information Criterion **(BIC)**, Adjusted BIC **(aBIC)**, and consistent Akaike Information Criterion **(cAIC)** [8]
- **Entropy** values near 1 to indicate better latent class separation [9]
- **Interpretability** which includes homogeneity (response probabilities close to 0 or 1) and latent class separation

**Table 2.** Latent Class Analysis Fit Indices with 1-6 Latent Classes

| Number of classes | Log likelihood | Resid. df | BIC | aBIC | cAIC | Likelihood Ratio | Entropy |
|---|---|---|---|---|---|---|---|
| 1 | -1372.480 | 255 | 2795.143 | 2766.609 | 2804.143 | 583.7178 | - |
| 2 | -1272.315 | 245 | 2650.573 | 2590.333 | 2669.573 | 383.3880 | 0.744 |
| 3 | -1248.654 | 235 | 2659.011 | 2567.066 | 2688.011 | 336.0665 | 0.667 |
| 4 | -1235.968 | 225 | 2689.398 | 2565.749 | 2728.398 | 310.6946 | 0.664 |
| 5 | -1225.375 | 215 | 2723.971 | 2568.616 | 2772.971 | 289.5079 | 0.721 |
| 6 | -1215.075 | 205 | 2759.132 | 2572.072 | 2818.132 | 268.9089 | 0.762 |

Results indicate that a **two-class latent model** is the best fit. We prioritize the first criteria of small BIC, aBIC, and cAIC values. While Class 3 provided a small aBIC value, the entropy value indicates that Class 2 has a better class separation. For interpretability, having two classes for coding attitudes can align with the EVT framework.

### Two-Class Latent Model

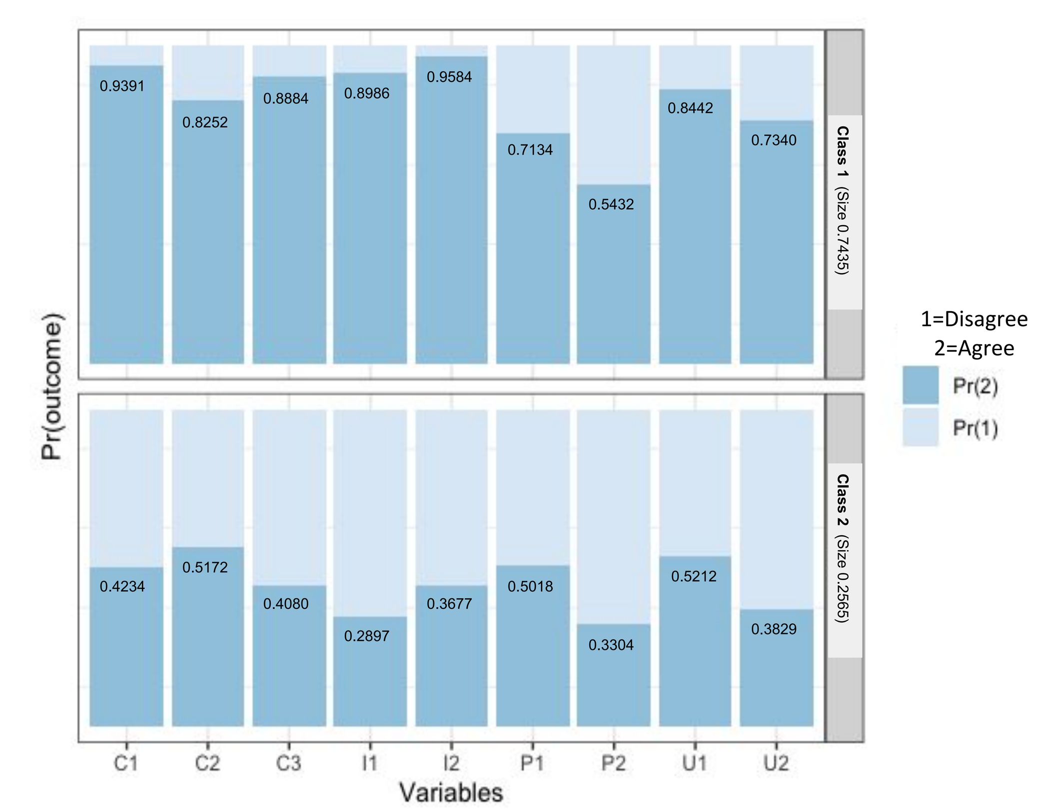
Class 1:
**High Success Exp & Value (H-SEV)**

The estimated class size is 0.7435 of the sample. 105 are female and 96 are male.

Class 2:
**Developing Success Exp & Value (D-SEV)**

The estimated class size is 0.2565 of the sample. 32 are female and 31 are male.

**Figure 2.** Two-Class Item Probability Bar Graph




## RQ2: Are social influences (i.e. family & friends) and students' sex predictors of coding attitude types?

**Table 3.** Logit Coefficients, Standard Errors (S.E.), t values, and p-values For the Two-class Solution with Sex (coded 2=female, 1=male), and Social Influences (S1, S2, S3). Using *HIgh Success Exp & Value* as the Reference Class.

| Covariate | Coefficient | S.E. | t | p-value |
|---|---|---|---|---|
| *Developing Success Exp & Value* | | | | |
| Sex | 0.33 | 0.40 | 0.83 | 0.405 |
| S1: My friends think coding is cool | -1.42*** | 0.40 | -3.61 | 0.000 |
| S2: My parents think coding is cool | -1.53*** | 0.41 | -3.74 | 0.000 |
| S3: I am friends with kids who code | -0.42 | 0.42 | -1.00 | 0.316 |

***p≤ .001, **p≤ .01, *p≤ .05

- There is **no significant difference** when comparing D-SEV to the H-SEV class with respect to **sex** (0.33, p> 0.05) and whether students **know friends who code** (-0.42, p> 0.05).
- Students have **friends who think coding is cool** (S1) are **significantly more likely** to be in the **H-SEV** class compared to the D-SEV class.
- Students who have **parents who think coding is cool (S3)** are are **significantly more likely** to be in the **H-SEV** class compared to the D-SEV class.

## References

(1) Code.org, CSTA, & ECEP Alliance. (2020). *2020 State of Computer Science Education: Illuminating Disparities*. Retrieved from https://advocacy.code.org/stateofcs
(2) Hogg, M. A., & Vaughan, G. M. (2011). *Self and identity.* Social psychology, 6, 111-146.
(3) Hallajow, N. (2018). *Identity and attitude: Eternal conflict or harmonious coexistence*. Journal of Social Sciences, 14(1), 43-54.
(4) Wigfield, A., Eccles, J. S., Fredricks, J. A., Simpkins, S., Roeser, R. W., & Schiefele, U. (2015). *Development of achievement motivation and engagement*. Handbook of child psychology and developmental science, 1-44.
(5) Education Data Partnership. (2019). *District Summary: Santa Ana Unified.* Retrieved October 15, 2020, from http://www.ed-data.org/district/Orange/Santa-Ana-Unified
(6) Mason, S. L., & Rich, P. J. (2020). *Development and analysis of the Elementary Student Coding Attitudes Survey*. Computers & Education, 103898.
(7) Linzer, D.A. and Lewis, J. (2013). *poLCA: Polytomous Variable Latent Class Analysis*. R package version 1.4. http://dlinzer.github.com/poLCA.
(8) 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.
(9) Celeux, G., & Soromenho, G. (1996). *An entropy criterion for assessing the number of clusters in a mixture model*. Journal of Classification, 13, 195-212.
(10) Kong, S. C., Chiu, M. M., & Lai, M. (2018). *A study of primary school students' interest, collaboration attitude, and programming empowerment in computational thinking education.* Computers & Education, 127, 178–189.
(11) Dou, R., Hazari, Z., Dabney, K., Sonnert, G., & Sadler, P. (2019). Early informal STEM experiences and STEM identity: The importance of talking science. Science Education, 103(3), 623–637.

## Discussion

- The insignicant finding for sex may be an indicator that this age-group have not developed sex biases in CS. There are mixed findings pertaining to sex differences in early CS attitudes. While previous work have resulted in an insignificant finding, a recent study revealed that boys had greater confidence and interest in CS/STEM among 4th-6th graders.[10]
- The insignificant finding of S3 reveals that access to coding peers does little to inform attitude types. Rather, the significance of S1 and S2 shows the value of interactions that intentionally cultivate coding interest.
  - The phrase for S1 and S3 pertains to a *perceived* "coolness" or what students deem as expressed interest from parents and friends. This level of engagement within social circles should be further investigated to determine what kinds of discussions are most impactful to coding attitudes.

## Implications

**Students are forming distinct coding attitude types by fourth grade despite a lack of formal CS education in elementary school**

- This study was able to include social influences of coding attitudes, but there is clearly more that needs to be identified so as to better inform the D-SEV class. This includes but is not limited to media content, access to technology, and parent education or profession.

**Importance of social interactions to develop CS attitudes**

- The findings align with prior literature on the critical role that social influences, especially parents, have on pre-adolescent identities in STEM.[11] This study confirms that computer science is not an exception.

**The need for early intervention of computer science learning**

- Currently, only four states have established policies to provide a form of CS primary instruction by the year 2021.[1]
- There is a clear group that has been exposed to social factors that support H-SEV in CS and another that has not. CS intervention in elementary school should thus actively address coding attitudes.

## Limitations & Future Analysis

•These findings can inform public school districts with similar demographics and computer science opportunities.

•The ESCAS survey was primarily designed to assess in-school coding interventions and overlooks other socio-cognitive factors that can influence coding attitudes.[6]

•This analysis used the R poLCA package which could not calculate the p-values of a bootstrap likelihood ratio test (BLRT) and Lo-Mendell-Rubin LRT test (LMR LRT). These are the appropriate methods to compare across nested latent class models unlike a traditional Likelihood Ratio Test.[8]

## Acknowledgements

I would like to thank my co-advisor Mark Warschauer and the IMPACT-CONECTAR team.

This material is based upon work supported by the National Science Foundation under Grant No. 1923136 and the US Department of Education under Grant No. U411C190092.