Studies in Educational Evaluation 70 (2021) 101025

Contents lists available at ScienceDirect

Studies in Educational Evaluation

journal homepage: www.elsevier.com/locate/stueduc

Co-enrollment density predicts engineering students' persistence and graduation: College networks and logistic regression analysis

Eric Leonardo Huerta-Manzanilla [a,*], Matthew W. Ohland [b], Rebeca del Rocío Peniche-Vera [c]

[a] Graduate School of Engineering, Universidad Autónoma de Querétaro, Edificio G de Posgrado Oficina 16, Centro Universitario Cerro de las Campanas s/n, Col. Las Campanas, Querétaro, 76010, Mexico

[b] School of Engineering Education, Purdue University, Armstrong Hall, 701 West Stadium Av., West Lafayette, IN, 47907-2045, United States

[c] School of Engineering, Universidad Autónoma de Querétaro, Parque Biotecnológico 3er Piso, Centro Universitario Cerro de las Campanas s/n, Col. Las Campanas, Querétaro, 76010, Mexico

ARTICLE INFO

Keywords:
Students persistence
Logistic regression
Engineering education
Receiver operating characteristic curves
Network analysis

ABSTRACT

College retention is a concern for educational institutions and researchers. This concern is particularly acute in engineering for reasons including workforce shortages, economic competitiveness, social justice, and socioeconomic equity. This study presents the evaluation of co-enrollment density (CeD) for engineering students at eight medium and large American public universities over 24 years. CeD is a novel metric estimated using enrollment records that may predict retention in 4-year bachelor of science programs in engineering. Graduation and persistence were fitted to CeD with logistic regression. Students in denser co-enrollment clusters—high CeD—tend to graduate more than their classmates in less dense neighborhoods—low CeD. The regression models predict graduation with odds ratio intervals 95 % CIs [3.24, 4.81] and area under the receiver operating curve [0.76, 0.80]. CeD is more sensitive to students who do not persist, particularly after the first year, so CeD's cut-off points may be indicators for dropouts' risk.

1. Introduction

There is evidence of declining interest in engineering and STEM careers (Becker, 2010; Belser et al., 2018; Johnson, 2013; Johnson & Jones, 2006; Sithole et al., 2017). It raises concerns over a workforce shortage and a concomitant negative effect on economic competitiveness (Committee on Prospering in the Global Economy of the 21st Century: An Agenda for American Science and Technology, 2007). Approximately 50 % of students who enroll in engineering graduate from those programs at some point (Aljohani, 2016a; Braxton et al., 2013). A small fraction (<20 %) graduate in four years, and less than half within six years (Johnson, 2013; Zhang et al., 2004). Retention is the rate at which students who start college graduate at the same Institution (Berger et al., 2012; Titus, 2004). The graduation rate for bachelor's degree students was 63.2 % at public 4-year institutions from 2003 to 2009 (Johnson, 2013). Student departure impacts educational success markers, yet institutions have not been able to mitigate it (Aljohani, 2016a; Braxton et al., 2013; Tinto, 2006). This study is motivated by the need to get metrics to forecast students at risk for dropout—early and efficiently (Sithole et al., 2017).

1.1. Definitions

The terms related to students leaving college are not standardized. Table 1 shows the concepts used as a framework for this work, and Fig. 1 shows their primary relationships. The definitions are not meant to oversimplify the complexity of the phenomenon (Rintala et al., 2012) but provide a conceptual framework for this work. Persistence and retention are the student and institution's perspectives of continued enrollment; similarly, dropout and attrition are perspectives on students leaving college. Attrition does not imply that the student does not ever graduate, just that he/she dropout from a particular institution.

1.2. Related work

Retention is a critical issue in tertiary education; therefore, it has been studied from many different perspectives; however, the diversity of the educational settings where the phenomenon occurs, and its campus-based nature, prevents generalizable results (Aljohani, 2016b; Barbera et al., 2020). The development of multi-institutional datasets (Ginder, 2013), such as the one used in this research (Ohland & Long, 2016), has

* Corresponding author.
E-mail addresses: ehuerta@uaq.edu.mx (E.L. Huerta-Manzanilla), ohland@purdue.edu (M.W. Ohland), peniche@uaq.mx (R.R. Peniche-Vera).

https://doi.org/10.1016/j.stueduc.2021.101025
Received 11 October 2020; Received in revised form 8 April 2021; Accepted 25 April 2021
Available online 3 May 2021
0191-491X/

Table 1
Definitions for terms related to enrollment and persistence in the context of the study.

Term	Definition
Attrition	A reduction in a school's student population because of transfers or dropouts;
Co-Enrollment	Students that voluntarily enroll in the same section of a course in the same term (CE);
Dropout	The temporary or permanent voluntary withdrawal from an education or training program before completion;
Dual-Enrollment	Enrollment of students in two schools at the same time;
Cohort	Group members that share a common educational experience. In the study context, students enrolled in the same courses and terms due to compulsory institutional practices.
Enrollment	The total number of individuals registered in a program accounts for a relationship between student and institution;
Graduation	Receiving a diploma or degree for completing a phase of formal education. It is an institutional and an individual goal.
Persistence	The continuance of a student's enrollment from the first to the second year—measured as the enrollment in one additional term after the first year. It is measured by the percentage of students who return to college for their second year (NCES, 2019). In this study, retention rate and persistence rate are the same;
Retention	The ability of an educational institution to prevent student attrition and keep students enrolled until graduation. Its rate is measured as the percentage of students who return to the same institution (NCES, 2019). In this study, retention rate and persistence rate are the same;
Transfer	Students who have transferred or intend to transfer from one higher education institution to another to achieve more advanced or different educational goals (College transfer students).

Note. These definitions were adapted from the Educational Resource Information Center's Thesaurus (ERIC, 2020), except the terms that include a citation.

facilitated identifying general trends typical of multiple institutions and programs, and institution-specific findings.

The engineering curriculum is diverse (Corlu et al., 2018). The roots for western engineering education, as known today, may be traced to early European technical schools. Continental Europe approached engineering as a public service involving knowledge of advanced mathematics and science; the École Nationale des Ponts et Chaussés in France is an example of such schools. In contrast, Anglo-American engineers were trained on the job. England's early engineering schools represented such a model that evolved after World War I when industries demanded engineers with higher scientific knowledge levels (Corlu et al., 2018). However, even today, there are many skills and knowledge to train engineers (Lucena et al., 2008; Passow & Passow, 2017). The diversity of the discipline's curriculum offers an additional layer of complexity to the inquiry on persistence and graduation in engineering, which is part of the emerging field of research in engineering education (Borrego & Bernhard, 2011). Jesiek et al. (2009) observed a sense of ambiguity about the identity and status of engineering education research.

Problem/Project–Based–Learning (PBL) and Conceive–Design–Implement–Operate (CDIO) are initiatives that propose to reform the engineering curriculum and may impact engineering education at the colleges involved in them (Bennedsen et al., 2019; Edström & Kolmos, 2012, 2014; Malmqvist et al., 2015). Regional projects like the ATTRACT (Enhance The Attractiveness of Studies in Science and Technology) may improve STEM degrees in Europe. ATTRACT's preliminary results were reported for the ATTRACT's work package-8 "Students Retention" that ran from January 2010 to October 2012 (Rintala et al., 2012). The actions to improve retention were divided into three strategies: Structure of studies, progression rules, and human support. Universities in partnership with the project reported that the levels of student's preparation, commitment, and motivation, along with study skills, had the highest impact on attrition (Rintala et al., 2021). The American Board of Engineering and Technology's outcomes-focused criteria for engineering program accreditation are examples of a quality-driven approach for designing and implementing the engineering curriculum (Akera et al., 2019; Dobryakova & Froumin, 2010). There is evidence that engineering program accreditation improves retention (Al Busaidi, 2020). The evolving nature of engineering education makes the study of retention a challenge, but it should not be an excuse to stop the efforts on its research and improvement.

According to Aljohani (2016a), "The larger body of student retention studies were designed and conducted in the American higher education contexts." Nevertheless, there are notable empirical studies and strategies to improve retention in Australia, Great Britain, and Europe. Australian higher education institutions and government have developed projects to improve retention focused in the first year. They have approached the retention issue based on educational experiences' quality (Hodges et al., 2013; Krause & Armitage, 2014; Willcoxson et al., 2011). The University-Experience Survey, the Course-Experience-Questionnaire, and the First-Year-Questionnaire are examples of instruments to understand how the quality of education impacts its outcomes. Hodges et al. (2013) reported that programs providing free educational access to government-targeted equity groups had higher attrition rates than conventional undergraduate degree programs, suggesting that the lack of financial interest leads to less accountability and engagement. The argument was in line with student engagement theories (Tight, 2020). Jones (2008) reviewed ten fundamental studies on retention in the British educational system. The attrition factors were classified in individuals' characteristics, such as educational goals, preparation for college education, abilities, institutions' teaching quality, and fit and satisfaction with the institution.

There are more than four decades of literature on retention (Aljohani, 2016b; Burke, 2019; Melguizo, 2011; Tinto, 2006). The study of retention/persistence has evolved from a psychological approach (student's attributes) to a sociological perspective focusing on the student/college relationship (Astin, 1999; Lin, 2020; Tinto, 1997, 2006). Despite the large and more recent body of research on retention, the *Longitudinal Model of Student Departure* (Tinto, 1975, 1993) still has paradigmatic status. While Tinto's theory has enlightened the subject's

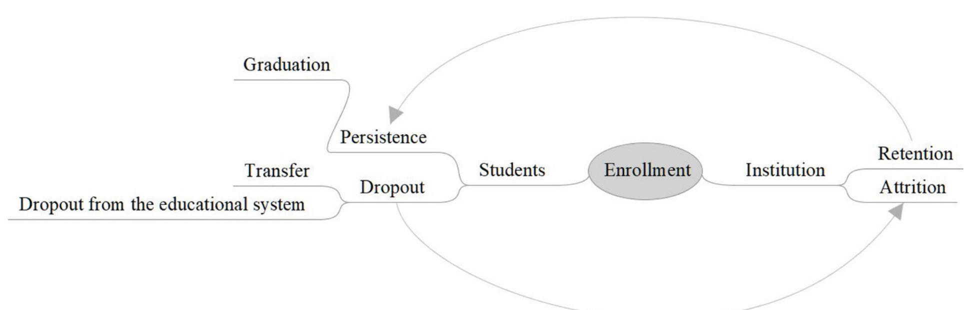


Fig. 1. Concepts related to graduation and their relations.
Note. The diagram shows the direct relations between the concepts on students leaving college (see their definitions in Table 1), from two perspectives: the institutional at the right, and the students' at the left.

inquiry, it presents limitations: It applies to traditional, non-minority, homogeneous student-bodies enrolled in 4-year programs, and it does not consider external factors. Further, Tinto did not provide empirical evidence on the relation of integration and persistence or even an instrument to measure integration (Braxton et al., 2000, 2013; Melguizo, 2011; Seidman, 2005). This paradigm has made the longitudinal path analysis expected to inquire on retention (Chapin, 2019; Cominole et al., 2007; Lee & Ferrare, 2019; Ohland & Long, 2016; Wang & Wickersham, 2014; Zhang et al., 2004). Tinto (1997) recognized some of the limitations of his theory, proposed a different perspective based on learning communities, and remarked:

> "…we would be well served by supplementing our use of path analysis to study the process of persistence with *network analysis and social mapping of student interaction patterns*. These will better illuminate the complexity of student involvements and the linkages that arise over time between the classroom and out-of-the-class experiences."—the emphasis is ours.

The study of college student networks is still an emerging field, with few examples exploring the classroom/course relations, and even fewer relating it with persistence (Biancani & McFarland, 2013; Forsman et al., 2014, 2015). Yet, the study of college student networks may assist path analysis to uncover the complexity of students' interaction patterns and their impact on persistence. Network analysis typically requires the acquisition of relational data, which is usually obtained with surveys. Thomas (2000) published a study of records and surveys for 322 first-year students, applied a network analysis, and fitted the results to persistence constructs. The author found that GPA, intent, institution commitment, and goal commitment explained 26 % of the persistence variance; each of these factors had a more significant predictive relationship than the two measures of integration combined. Grunspan et al. (2014) published a classroom network analysis; the authors related the network's positions with success on exams. The data collection was survey-based, and the construction of relational data applied formal network analysis.

Gardner et al. (2018) and collaborators probed that a CE network may predict grades; this work was the closest to ours. Israel et al. (2020) reported the analysis of a student–course network at the University of Michigan. They found that students of the same majors clustered, but transfer students showed low centrality. These results were expected; however, the authors did not relate the network's parameters with academic outcomes. Their assumption that sharing a lecture hall implies social contact was unwarranted. Some studies use the term co-enrollment to refer to dual-enrollment, which is defined earlier and is conceptually different (Crisp, 2013; Wang & McCready, 2013; Wang & Wickersham, 2014). CE does not provide the type of relational data required for orthodox social network analysis, which may explained why there are so few examples of such an application. Instead of assuming that CE implies social contacts, we propose that CE may be a proxy for academic effort levels and that CeD may reveal clusters of students with similar commitment with the goal of graduation. The Data and Methods section shows the network analysis that explores this idea (Kang, 2019; Niu, 2020).

2. Data and methods

Sharing a lecture hall does not imply social connection; instead, we proposed that $CeD = \log_2(CE)$. CE density in a network (see Fig. 2) reveals how the students clustered. We explored how the clusters relate to the probability of graduation. CeD represents similarities in the student's effort in navigating the curriculum and how committed they are to academic goals (Braxton et al., 2013; Nicoletti, 2019; Tinto, 1993, 1997). Students who enroll in the same courses simultaneously require similar perception and attaintment of the curriculum, intellectual development, and motivation. Students that not co-enroll differ in academic achievement levels. Therefore CeD reveals clustering patterns that may be related to student's engagement and interaction with peers and faculty, that are known to be connected with persistence (Astin, 1999; Braxton et al., 2000, 2013; Cabrera et al., 1992; Tinto, 2006); particularly at the first year of studies (Braxton et al., 2013; Cabrera et al., 1992; Nicoletti, 2019; Tinto, 1993) (Table 2).

2.1. Description of the data

We used the Multiple-Institution Database for Investigating Engineering Longitudinal Development, or MIDFIELD (Ohland & Long, 2016). It contains whole population records for a sample of institutions that offer undergraduate engineering programs. No institution on record was known to have cohorting (definition in Table 1) practices in its programs. A general lack of cohorting arises based on variability in mathematics readiness (Blat, 2018) and other preparatory coursework, differentiation in student educational goals, and multiple sections due to class sizes that cannot meet the total course enrollment needs. Table 3 shows a summary of graduation rates for the eleven public universities in MIDFIELD at the time of the study. Three institutions (A, G, and K) could not be included in this study due to the lack of section data required to differentiate separate offerings of a course given in the same term. Detailed descriptions of the institutions are not provided to protect their confidentiality. Still, the MIDFIELD database used in this work was shown to be representative of a national database of engineering programs to the extent that such a comparison could be conducted (Orr

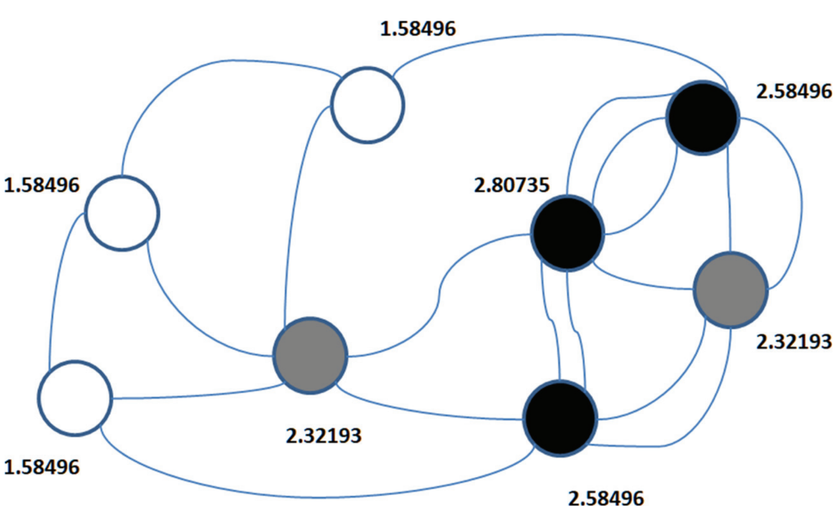


Fig. 2. A small student-course network example.
Note. The numbers are CeD values, the arcs are the courses co-enrolled, and the nodes are the students. Darker nodes represent students in a high CeD cluster and lighter nodes students in low CeD clusters. Darker nodes represent students with similar effort levels in navigating the curriculum and may have identical involvement patterns.

Table 2
Notations.

ω	Coenrollment (CE) in student's networks;
X: Ω	A random variable in a sample probability space
$AA^T = [x_{ij}]$	Adjacency matrix where A is the affinity matrix.
$x_{ij} \in AA^T$	Non-diagonal elements in the adjacency matrix;
k	A constant;
$x_{i=J}$	Adjacency matrix's diagonal element;
ψ	Coenrollment density (CeD), or neighborhood clustering;
Y	Graduation (4, 6 Yr., ever), 1 Yr. Persistence: 1 True, 0 False;
ξ	Logistic model error, or model's residual;
β_i	Logistics regression coefficients;
$OR = p(Y)/(1 - p(Y))$	Odds ratio;
$\lambda = \log(OR)$	Logit: Logarithm of the odds ratio;
d	Cohen's D, a standard measure of effect size;
CI	Confidence interval, set at 95 %, unless otherwise specified;
CUTj	Cut-off values for CeD based on Youden's index (j);
AIC	Akaike's entropy-based information criterion;
ROC	Receiver operating characteristic curve;
AUC	The area under the receiver operating characteristic curve;
$Sen = TP/(TP + FN)$	Sensitivity, true fraction TP=True positives, FN=False negatives;
$Spe = TN/(TN + FP)$	Specificity, true negative TN=True negatives, FP=False positives;
CeD1,CeD2,CeD3, CeD4	CeD based on 1, 2, 3, and 4-year networks.
P1, G4, G6, EG	P1-First-year persistence, G4, G6, and EG graduation at 4, 6, and ever;

Table 3
Descriptive statistics for institutions in the MIDFIELD database at the time it was available for the study. Institutional population ranging from 44,048 to 149,407. The population at each Institution cannot be shown to prevent institutional identification.

Institution	Graduated		Total
	No Grad.	Grad	
A	48.8 %	51.2 %	100 %
B	67.9 %	32.1 %	100 %
C	48.7 %	51.3 %	100 %
D	42.6 %	57.4 %	100 %
E	66.3 %	33.7 %	100 %
F	45.8 %	54.2 %	100 %
G	46.6 %	53.4 %	100 %
H	60.8 %	39.2 %	100 %
I	45.5 %	54.5 %	100 %
J	43.5 %	56.5 %	100 %
K	63 %	37 %	100 %
Total	50.7 %	49.3 %	100 %

et al., 2020). After removing the three institutions without the necessary data, the database included records for 68,293 first-time-in-college (non-transfer) engineering students. MIDFIELD has been the focus of ground-breaking research, including papers selected as the best published in the Journal of Engineering Education in 2008 (Ohland et al., 2008), 2011 (Ohland et al., 2011), and 2019 (Lord et al., 2019). Table 4 shows the records for engineering students' first-year persistence (P1) and graduation rates (G4, G6, and EG). 71.8 % were first-year persisters, 36 % graduated in four years, 47.8 % in six years, and 52 % in total.

2.2. Methodology

2.2.1. The research questions

The research questions were: (1) If$Y_1 \approx f(\psi_1)$, where Y_1 is first-year persistence, and ψ_1 is CeD at first-year; (2) If$Y_4 \approx f(\psi_2)$, where Y_4 is graduation at four-years, and ψ_2 is CeD at second-year; (3) If$Y_6 \approx f(\psi_3)$, where Y_6 is graduation at six-years, and ψ_3 is CeD at the third-year. This work used 24 models to evaluate questions (1), 16 models to evaluate each of questions (2) and (3), in addition to other complementary models (See Fig. 3).

2.2.2. CE (ω) undirected graphs

Fig. 4 describes the analysis to get the affinity matrix (A) from enrollment records. Unique course identifiers included the course ID, a term's code, and a section identifier. In the example, c111 is course 1, offered in term 1, in Section 1. The AA^T student-course affinity matrix product computes the adjacency matrix, times its transpose, its columns and rows have the CE data. The sum of rows less the diagonal is the total undirected graph per student in Eq. (1).

$$\omega = \sum_{j \in J} (x_{ij}) - x_{i=J} \quad \forall x_{ij} \in AA^T \tag{1}$$

2.2.3. Computing CeD (ψ)

The enrollment record (See the example in Fig. 4) shows that the student identified as s1 accounted 1 graph to s2, and 2 graphs to s3, for a total of $\omega_{s1} = 1 + 2 = 3$, then its co-enrollment density is $\psi_{s1} = \log_2(3) = 1.585$. It estimates the total peer contacts, a proxy for mutuality (Rao & Bandyopadhyay, 1987). However, CeD reflects the student's course selections that met other students' selections, not the social interaction, i.e., students may join in a course due to their academic effort level, not because they reciprocate socially.

$$\psi = \log_2(\omega) \tag{2}$$

2.2.4. Testing the hypothesis

Predicting graduation requires identifying a function $f(s)$ to map the change in Y's odds ratio probability from negative to positive $OR = p(Y)/(1 - p(Y))$, with a potential predictor (ψ in this case). One solution is called the logit function. See Eq. (3).

Table 4
Persistence (P1) and graduation (G4, G6, and EG) data for engineering students.

	P1		G4		G6		EG	
Inst.	No Pers.	Persist	No Grd.	Grad.	No Grd.	Grad.	No Grd.	Grad.
B	31.5 %	68.5 %	80.5 %	19.5 %	70.2 %	29.8 %	67 %	33 %
C	40.1 %	59.9 %	79.3 %	20.7 %	70.7 %	29.3 %	67.5 %	32.5 %
D	28.4 %	71.6 %	60.2 %	39.8 %	48.3 %	51.7 %	43.4 %	56.6 %
E	29.4 %	70.6 %	58.3 %	41.7 %	52 %	48 %	50.2 %	49.8 %
F	25.7 %	74.3 %	59.8 %	40.2 %	50.8 %	49.2 %	44.7 %	55.3 %
H	58.4 %	41.6 %	80.2 %	19.8 %	76.9 %	23.1 %	76.4 %	23.6 %
I	21 %	79 %	63.7 %	36.3 %	45.7 %	54.3 %	43.9 %	56.1 %
J	27.5 %	72.5 %	70.9 %	29.1 %	55.2 %	44.8 %	51.4 %	48.6 %
Total	19238	49055	43741	24552	35653	32640	32809	35484
	28.2 %	71.8 %	64 %	36 %	52.2 %	47.8 %	48 %	52 %

Note. P1 is first-year persistence; G4 and G6 are graduations at four and 6-year, and EG is ever graduated.

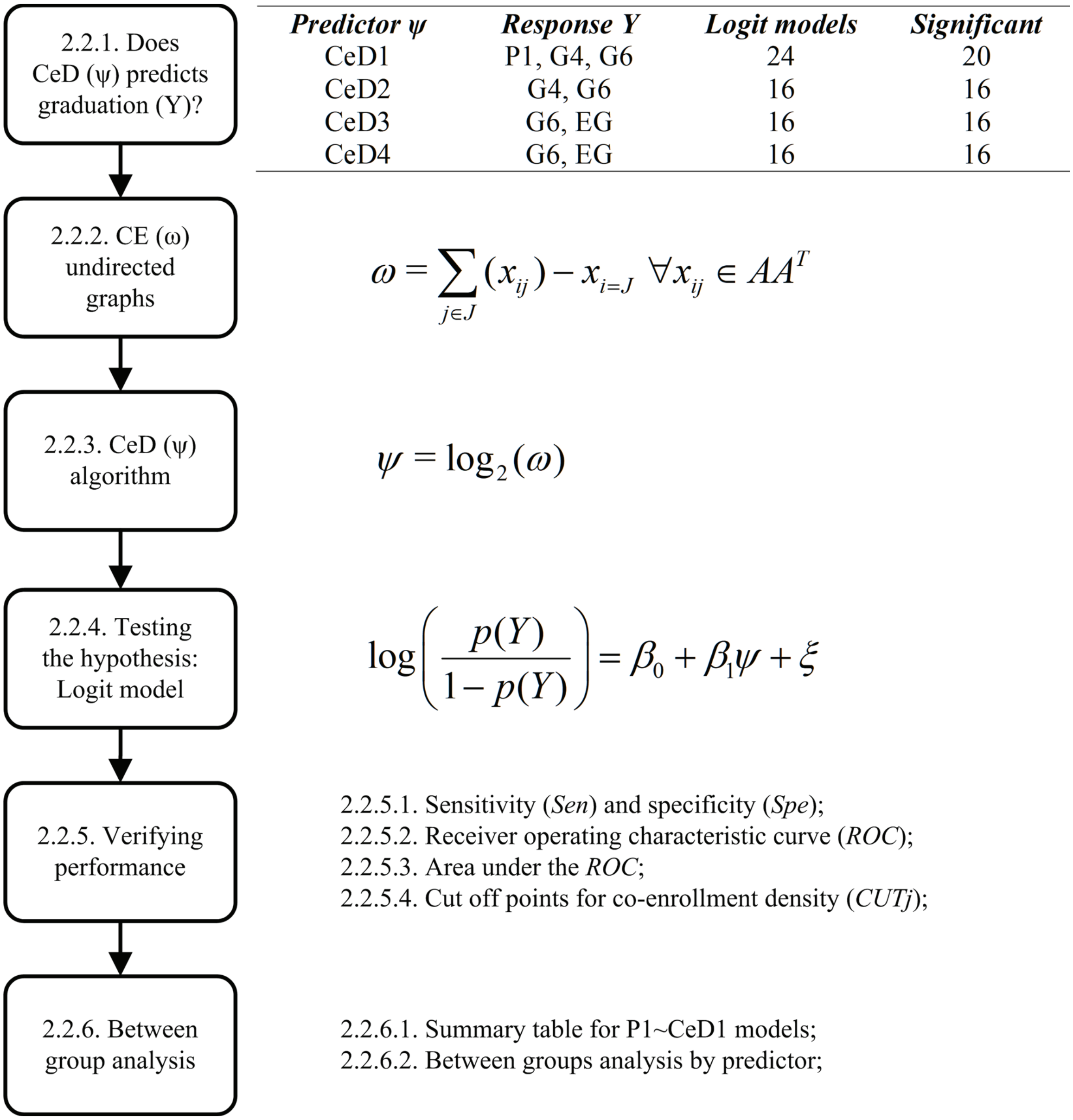


Fig. 3. Methodology overview.

$$p(Y) = \int_{-\infty}^{\psi} f(s)ds = \frac{\exp(\beta_0 + \beta_1\psi)}{1 + \exp(\beta_0 + \beta_1\psi)} \tag{3}$$

Where $f(s) \in X : \Omega$ and integrates the link function in Eq. (4), the logistic regression model (Dobson & Barnett, 2008, p. 126). CeD is the predictor for the logit of Y (graduation/persistence).

$$\log\left(\frac{p(Y)}{1 - p(Y)}\right) = \beta_0 + \beta_1\psi + \xi \tag{4}$$

Fig. 5(a) shows Y mapped to CeD3 predicting graduation at 6-year (G6) for Institution B. Its coefficient is the OR $\beta_1 \in (4.2, 5.79)$.

CeD was computed for college networks filtered for courses offered at the first, second, third, and fourth years of enrollment. CeD1 was tested to predict P1, G4, and G6. CeD2 was fitted to G4 and G6. CeD3 and CeD4 were fitted to G6 and EG. Nine models were obtained per Institution for a total of seventy-two logit models.

2.2.5. *Verifying the model's performance*

2.2.5.1. Sensitivity (Sen) and specificity (Spe). Sensitivity (*Sen*) is the proportion of true positives divided by the sum of true positives plus the false negatives (Eq. (5)), also known as true positive fraction, where *TP* are the true positives and *FN* are the false negatives predicted by the model.

$$Sen = \frac{TP}{TP + FN} \tag{5}$$

Specificity (*Spe*) is the true negative fraction in Eq. (6). TN is the true negative, and *FP* the false positive predicted by the model. The models obtained show a propensity to be better at *Spe* than at Sen, making them better to identify potential dropouts.

$$Spe = \frac{TN}{TN + FP} \tag{6}$$

2.2.5.2. Receiver operating characteristic curve (ROC). A receiver operating characteristic curve (*ROC*) is built with *Sen* and *Spe*. It verifies if

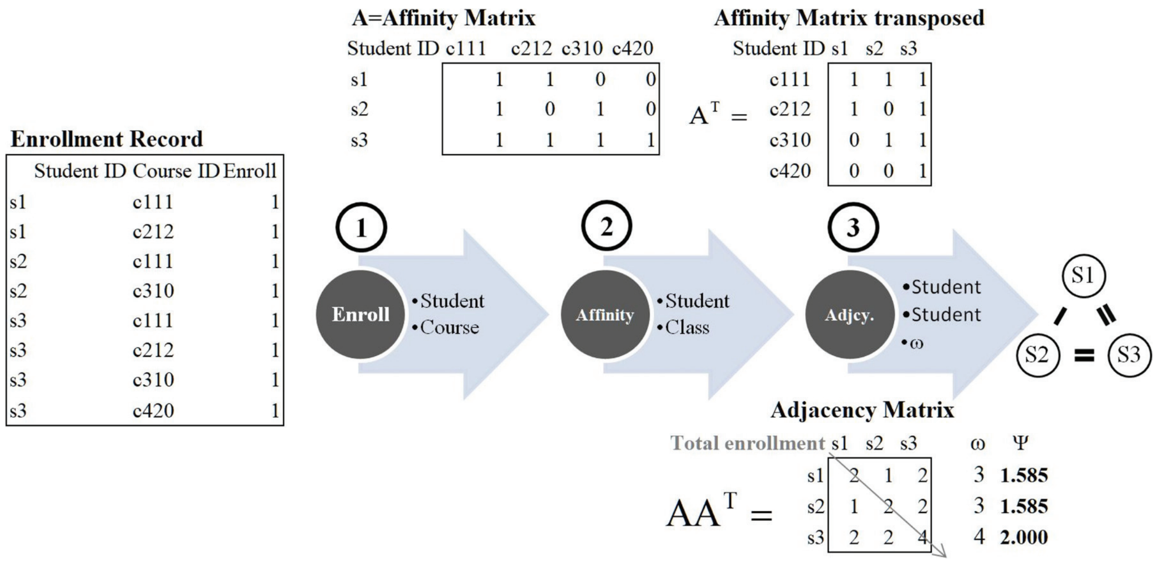
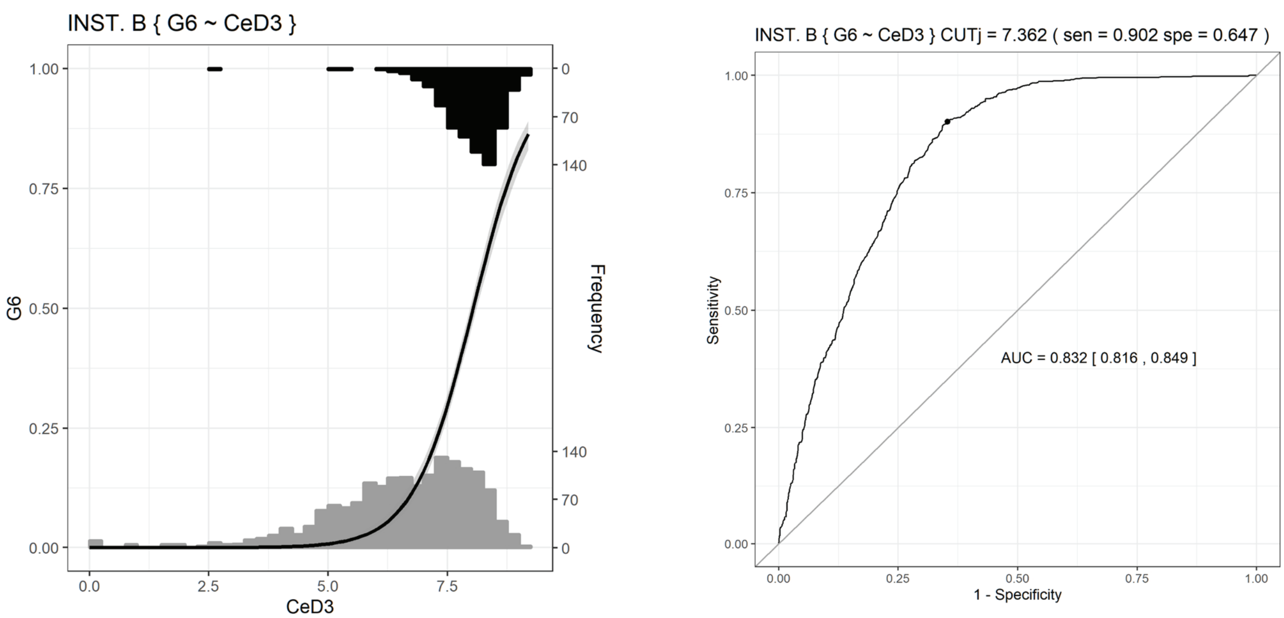


Fig. 4. Algorithm to compute CeD from enrollment records.



(a) Logistic regression curve.

(b) ROC Curve.

Fig. 5. Figure Logistic regression and ROC for Institution B: CeD@3 Yr predicting G6.

the logistic model correctly identifies the students who graduate (*Sen*) and the student who may fail to graduate (*Spe*). Fig. 6 shows two ROCs, and the diagonals represent a model that predicts the outcome no better than guessing ($AUC = 0.5$). Display 5(a) is a ROC for a model no better than guessing, and 5(b) is a model with good performance.

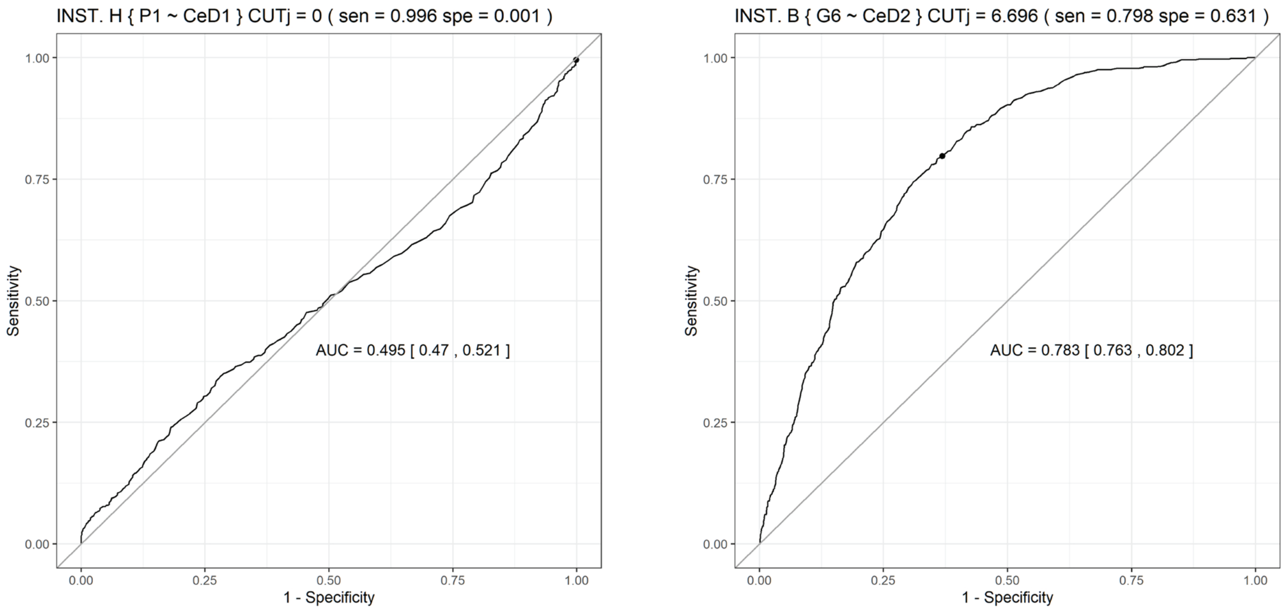
2.2.5.3. Area under the receiver operating characteristic curve (AUC). Fig. 6(a) shows the logistic regression of *CeD1* predicting first-year persistence (P1) at Institution H, a model with no prediction power (AUC∼.5). Panel 5(b) shows the *ROC* curve for Institution B with *CeD2* predicting *G6*, a good performance $AUC \in (.763, .802)$ model. *AUC* is a logistics regression metric that verifies its predictive performance for a binary classifier.

2.2.5.4. Cut off values (CUTj). Cut-off values (*CUTj*) can be estimated to optimize students' discrimination likely to graduate against those less likely to do it. CUTj was computed based on Youden's index (Fluss et al., 2005). Fig. 7 shows an example of the kernel smoothing for Sen + Spe, where $CUTj = 6.893$ is at Sen + Spe = .83 + .59 = 1.42. Fig. 8 shows the kernel smoothing and AUC for CeD1 predicting P1 at Institution E. CUTj = 7.2, Sen = .33, and Spe = .95, so this model predicts 95 % of students not enrolling one term after the first year.

There are other methods to estimate a *CUTj* value besides Youden's, given that an appropriate logistic regression model exists and its ROC curve has been computed. Each technique would produce a *CUTj* with slightly different Sen, and Spe fitted to an institution's policies.

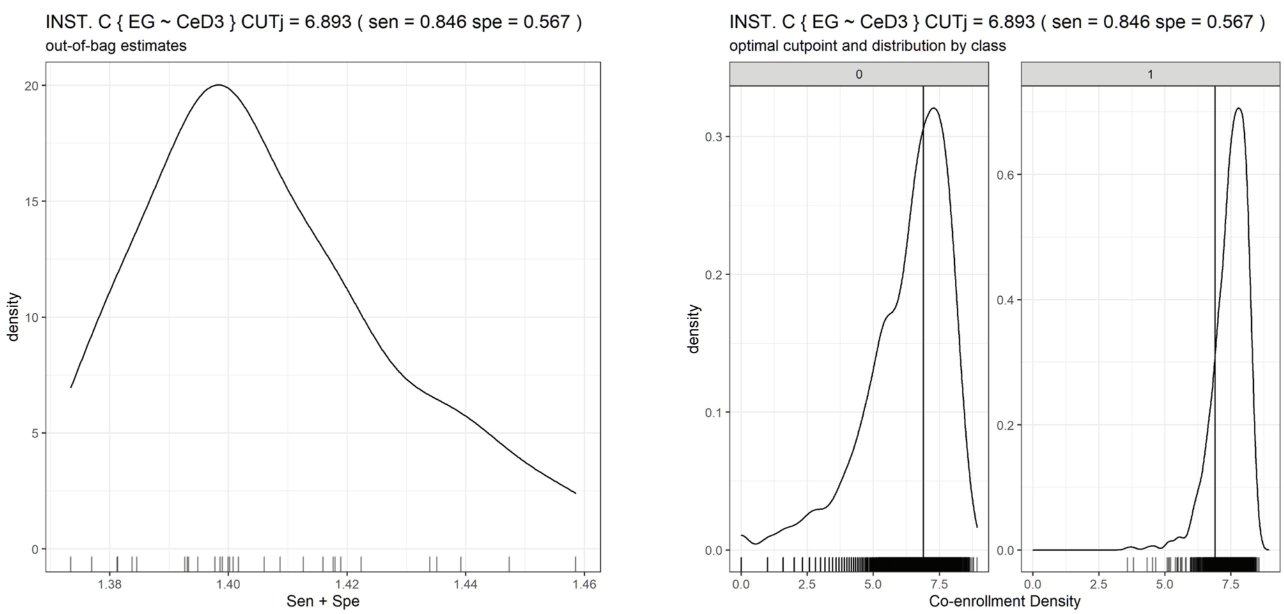
2.2.6. Between-group analysis

The logit model's parameters are reported in comparative analysis by the predictor, response, and institutions. Figs. 9–15present box plots for each model's odd ratios (OR), the area under the receiver operating



(a) *ROC* for CeD1 predicting P1. (b) *ROC* for CeD2 predicting G6.

Fig. 6. ROC curves: Left, Institution H. Right, Institution B.



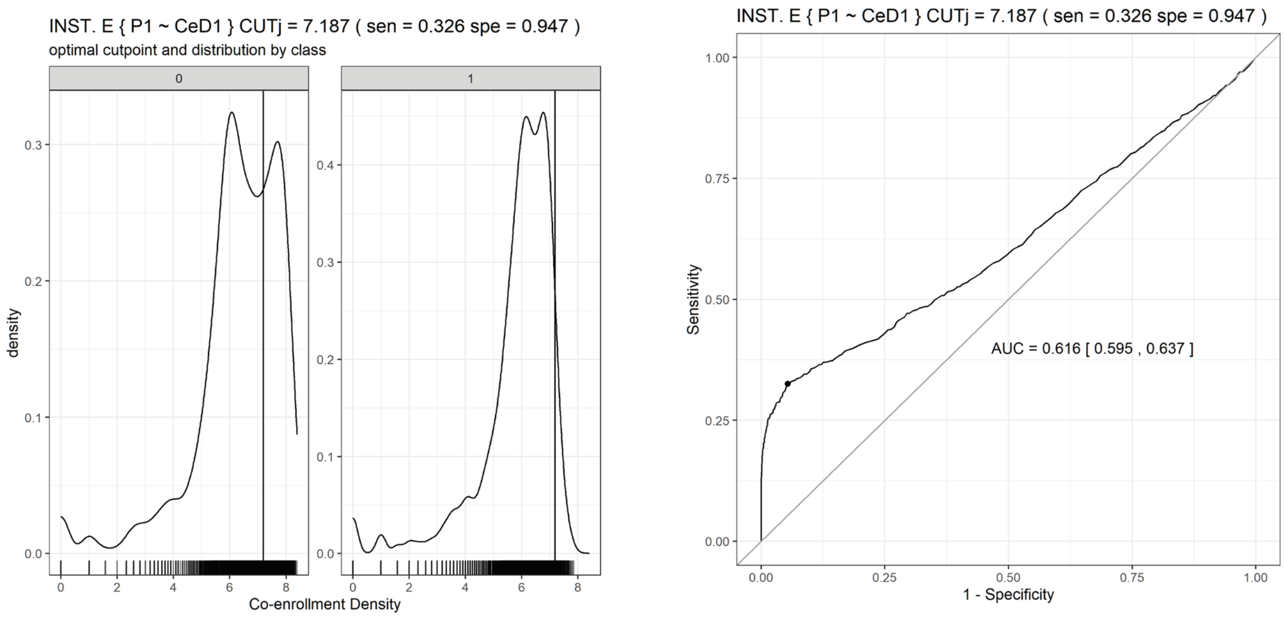
(a) *Kernel smoothing for Sen + Spe.* (b) Per class distribution.

Fig. 7. Institution C: CeD3 predicting EG.

curve (AUC), Specificity (Spe), and Sensitivity (Sen); all are aggregated by predictor and by the response. Cut-off points for CeD (CUTj) aggregated by predictor are presented in Fig. 15. The graphs feature box plots with standard quartiles and data points. Each box plot shows the median and its CI. At the top, there are brackets $p_{adjusted}$ displayed for pairs $p < .05$. The p-values were adjusted with Dunn's test (Dinno, 2015).

3. Results and discussion

Retention is a campus-based phenomenon; thus, empirical studies are usually of limited generalizability due to their descriptive nature based on particular samples, involving programs with specific designs and lengths (Aljohani, 2016b; Barbera et al., 2020). Multi-institutional data, especially spanning multiple enrollment years, address these limitations to some extent. Our study's contribution is that it includes over 20 years of data at eight institutions offering 4-year engineering bachelor's degrees, which are common in the United States and other regions. Kayis (2004) and collaborators, in a study of more than 100 engineering programs at more than 55 universities in North America, Canada, Europe, Asia, and Australia/New Zealand, reported that "the fourth year is the final year for most of the universities offering single degree engineering programs." Nevertheless, there are 3.5-year programs that lead to a technical engineering degree in Denmark, an example of a different undergraduate degree (Corlu et al., 2018).



(a) *CUT* and distribution by class. (b) *ROC* and *AUC*.

Fig. 8. Institution E: CeD1 predicting P1.

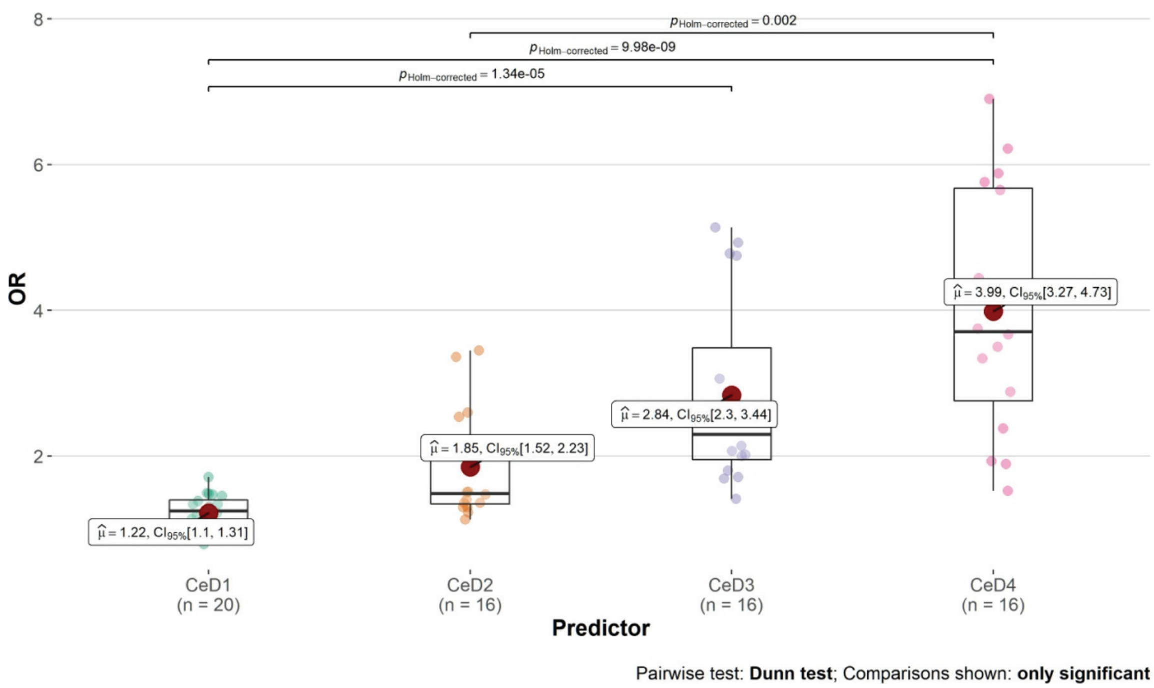


Fig. 9. Distribution of OR points across predictors (Significative Models).

CE data may neither predict graduation of transfer students nor explain students' causes of dropping from degrees or schools. It is only a proxy for other factors known to affect graduation's likeliness, such as involvement and commitment to the goal of graduation. Online and commuter programs may have different patterns in many aspects from residential ones (Braxton et al., 2013). The findings suggest that CeD describes student patterns related to persistence and graduation. There were 4/72 models for CeD1 that were not included in the comparative analysis because they did not meet one or more of the following criteria for significance: area under the ROC $.5 \in AUC(95\%CI)$, odd-ratio confidence interval $0 \in OR(95\%CI)$, or coefficients $p(Wald) > .01$. Table 5 shows the pairs of predictors and responses fitted in the 72 models that were computed.

3.1. The four non-significative models

The models: G4∼CeD1 at Institution I, G4∼CeD1, G6∼CeD1, and P1∼CeD1 at Institution H were not significant. The Institution I shows the lowest rate of first-year persisters in the sample, P1 = .416, well below the group average: P1 = .718. This condition—related to underlying factors at this Institution— may alter the students' enrollment patterns at an early stage notably.

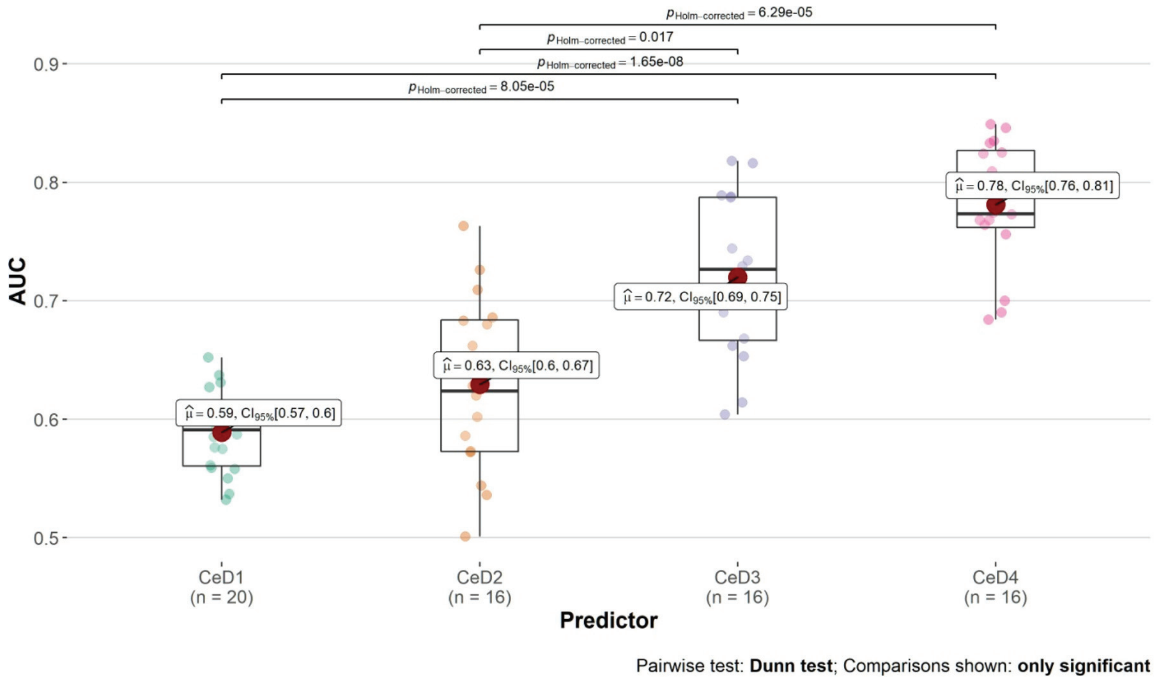


Fig. 10. Distribution of AUC across predictors (Significative Models).

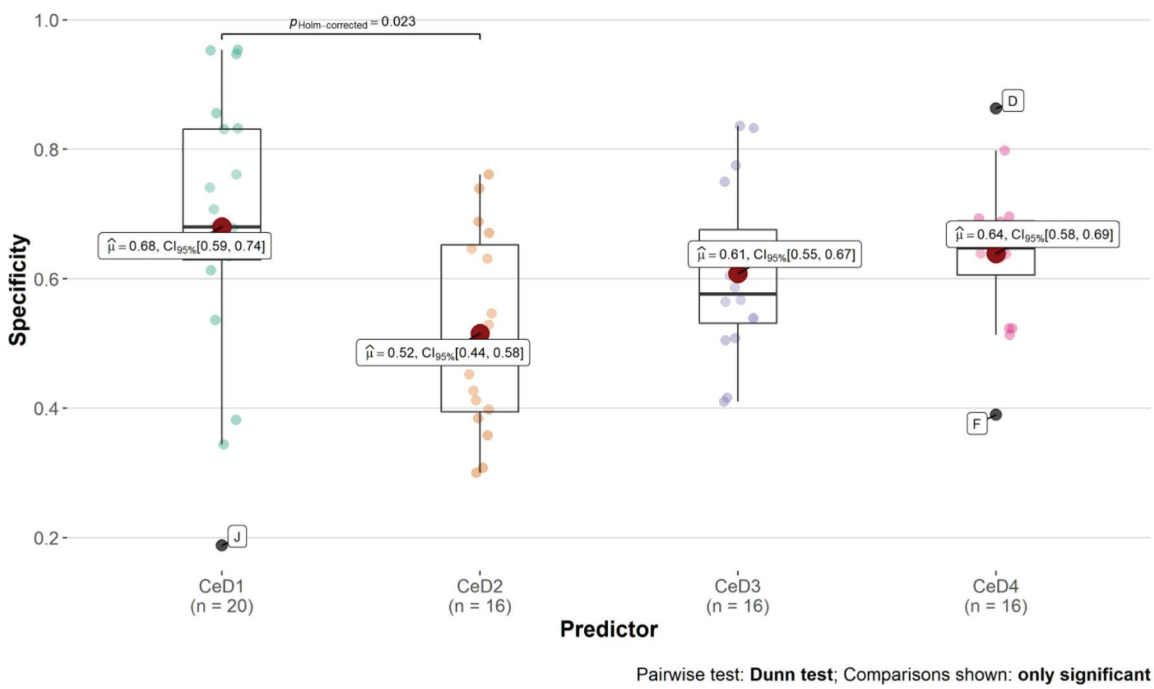


Fig. 11. Distribution of specificity points across Predictors (Significative Models).

3.2. CeD1 predicting first-year persistence

The models produced AUC > .5, except H's, which suggests that the models are better than guessing to predict P1. Institutions B to F's *CUTj* has *Spe*≥.678; that indicates that *CUTj* for CeD1 may correctly identify seven out of ten students at risk of leaving college. Thus, CeD1 is an effective first-year persistence indicator in 5/8 = 62.5 % of the institutions in the sample.

3.3. CeD predicting graduation

3.3.1. OR and AUC

Fig. 9 shows that CeD estimated for 1–4 years is related to graduation. The OR coefficients are more significant as the networks are evaluated with more courses in a curriculum. CeD may predict graduation since the first year. Fig. 10 shows the summary for AUC aggregated by the predictor. CeD estimated at the different years show a difference, except between CeD1 and CeD2. In general, we observe that CeD1 is a more helpful predictor than CeD2, particularly considering that CeD1 can be obtained one year before, and it also may predict P1. CeD at first

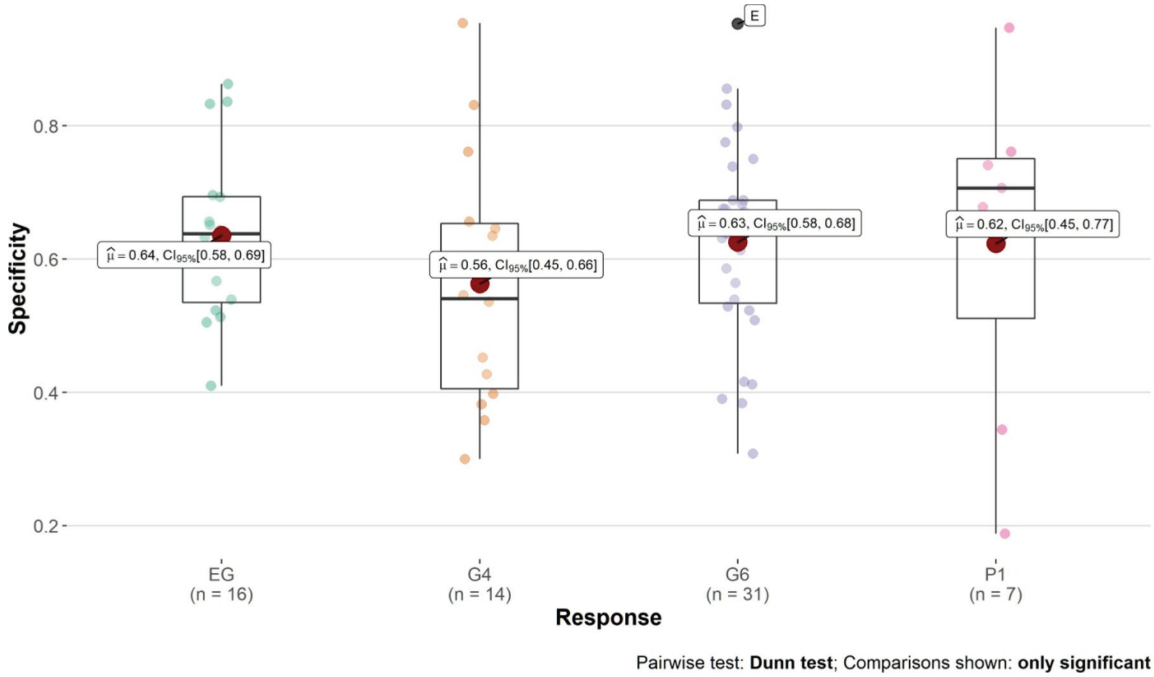


Fig. 12. Distribution of Specificity across Responses (Significative Models).

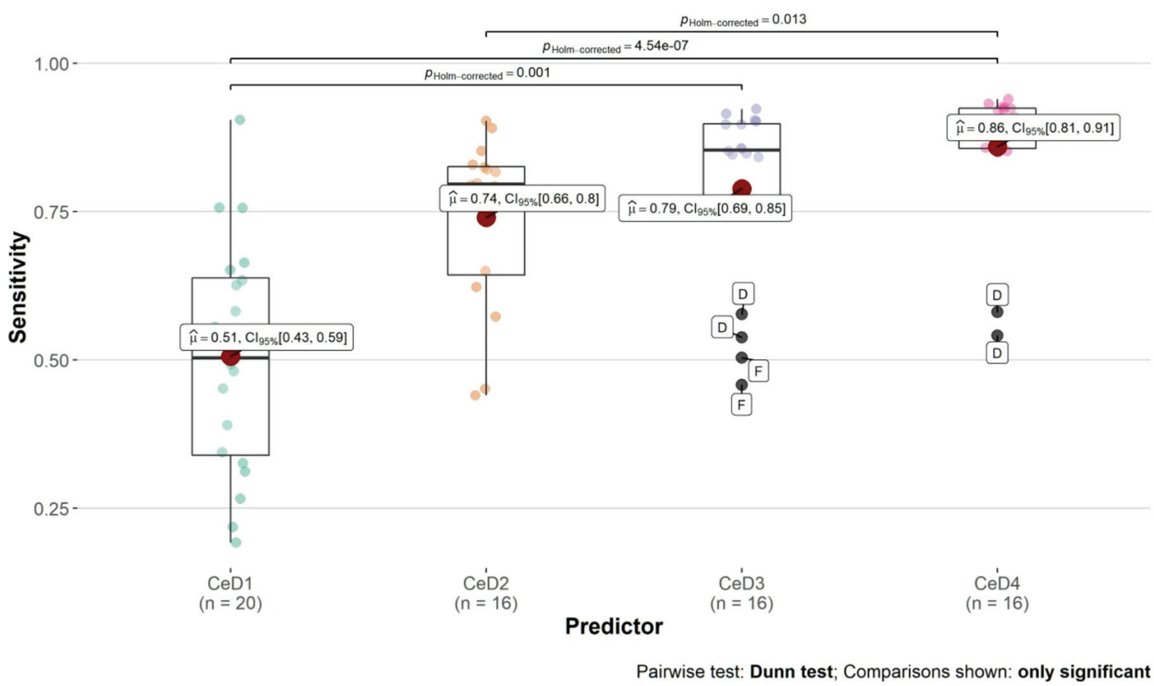


Fig. 13. Distribution of sensitivity points across predictors (Significative Models).

and the third years may be the most useful because they may be estimated earlier than CeD2 and CeD4, respectively.

3.3.2. Specificity

Figs. 11 and 12 present the summary and pairwise comparisons for Specificity by predictor and by responses. This set includes predictions for P1, G4, and G6. The models may identify students at risk of leaving college after the first year (CeD1), with specificities of Spe = .68 (Fig. 11). The chart shows that CeD at any year may predict the outcomes as they were fitted to CeD in the study. There is no significant difference between CeD3 and CeD4. Fig. 11 shows that CeD1's expected average Specificity is higher than CeD2's and similar to CeD3 and CeD4. This chart adds evidence to consider CeD1 as a good predictor for P1, G4, and G6 and suggests that it is a potential early warning index. Specificity estimated with logit models at three and four years shows lower variances. Particularly at four years. The models estimated at three and four years show similar results in terms of their expected average Spe.

Fig. 12 shows no significant difference between the achieved specificities to identify students at risk of leaving since the first year and then to not graduate at four, neither at six years. However, fourteen models (39.7 %) show Spe<.6. The other models (60.3 %) reached Spe≥.6. As

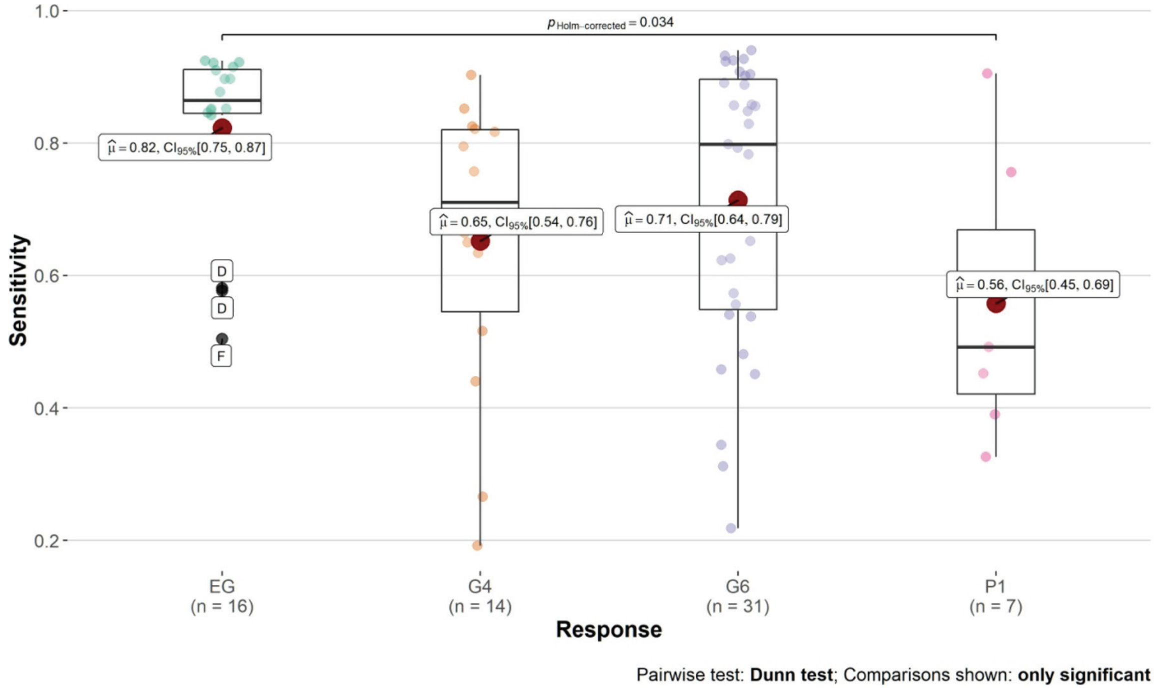


Fig. 14. Distribution of Sensitivity across Responses (Significative Models).

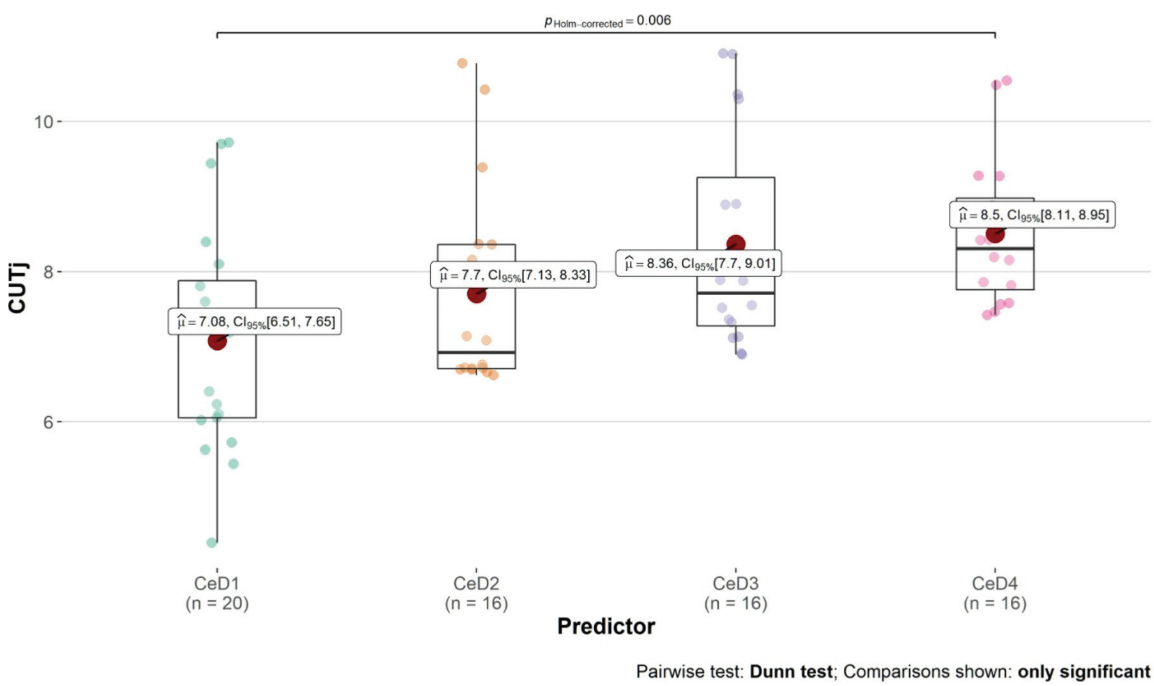


Fig. 15. CUTj by Predictor.

Table 5
The responses and predictors for the nine models per each Institution.

Predictor ψ	Response Y	Logit models	Significative
CeD1	P1, G4, G6	24	20
CeD2	G4, G6	16	16
CeD3	G6, EG	16	16
CeD4	G6, EG	16	16

mentioned earlier, the cut-off points that delivered these Spe levels could be estimated with different methods to more closely align them with institutional needs and policies.

3.3.3. *Sensitivity*

Figs. 13 and 14 show the results for Sensitivity aggregated by predictor and response, respectively. CeD1 offers a low capacity to identify those that may persist at the first year (P1) or graduate at four (G4) or six-year (G6). There is no significant difference for CeD2, but it is lower than those of CeD3 and CeD4. Institution D and F present outliers. CeD4 shows the highest Sensitivity of all the models.

The results provide evidence that supports the original thesis of the work. CeD is an index that may identify potential dropouts in the first year and those at risk of not graduating at the fourth, sixth, or ever. The evidence suggests that CeD is sensitive to students at risk (Specificity), and after the second year, it is also helpful to identify potential graduates (Sensitivity). Fig. 14 shows Sensitivity by responses. In this case, 69 % of the models reached Sen>.6. The prediction for the total graduates (EG) presents the most negligible variance; that implies that CeD3 and CeD4 were more robust predictors for graduation than the other CeD values computed with smaller networks, particularly P1.

3.4. Cut off points for CeD

CeD is a continuous predictor that allows the computation of cut-off points to identify potential dropouts and graduates. Fig. 15 shows no significant difference between the average estimates for cut-off points (*CUTj*) for the different outcomes investigated, and their variances were homogenous. This finding suggests that CeD and its *CUTj* is an intensive property for college networks, independent of their size.

4. Conclusions

The results suggest that CeD may be a robust and parsimonious predictor for first-year persistence and graduation at 4-year engineering programs. While MIDFIELD has been shown to be representative of a more comprehensive national database of engineering programs, it is impossible to determine if MIDFIELD is representative in this study because no other database exists capable of studying this phenomenon. Indeed, the United States Department of Education concluded that it was infeasible to create such a database (Cunningham & Milam, 2005). In particular, our findings suggest that institutions with extremely low persistence and graduation rates may be exceptions. CeD may replace multivariate models when formal enrollment records are available.

CeD allows the estimation of cut-off points that may help identify students at risk of not persisting after the first-year, or not graduating later. It shows higher Specificity than Sensitivity when estimated at one and two years. Therefore, CeD1 and CeD2 seem to be more sensitive to students at the risk of leaving. CeD3 and CeD4 are better to identify students that show a positive trend to graduation.

4.1. Implications for engineering programs and educational institutions

The method proposed may help institutions expand their efforts to identify students who could benefit from retention initiatives by including CeD or similar indexes. CeD1 provides a parsimonious index, and its cut-off threshold values may identify potential dropouts. It may be added to indicators known to affect persistence in the first year. The results may add a way to link the theory on persistence to engineering education practice using network analysis and empirical models based on course enrollment records. It is not known if CeD is related to institutional cohorting. Research has shown informal mentoring to be more effective than formal mentoring (Inzer & Crawford, 2005), which suggests that institutional cohorting (formal efforts to group students) may not have the same benefits or predictive value as found in this work.

4.2. Contribution to the literature on retention

The methodology and the CeD index add to the literature on retention. It is a system parameter that reveals patterns that the students tend to exhibit if they are more prone to stay in an educational program. CeD may be the first of other college network metrics that may support policies in improving college outcomes. Other researchers may use enrollment data to analyze college networks that may reveal student activity patterns related to academic outcomes. As Tinto (1997) suggested, network and data analysis may complement traditional longitudinal studies. CeD appears to be a novel and parsimonious metric that may predict retention for students in 4-year engineering degrees.

4.3. Future research

This work provides a framework to ask various research questions of value, as an extension of this work. Does CeD have similar performance in other disciplines and engineering programs with different designs? Since this study used aggregated data, its results are most applicable to most U.S. engineering programs. It would be valuable to ask how CeD and its usefulness vary based on an intersectional combination of race/ethnicity and sex. It would be helpful to study whether curriculum frameworks such as CDIO or PBL alter CeD and its predictive value. As suggested above, does institutional cohorting produce the same effect as spontaneous CeD? How does CeD perform in private colleges? Investigating the outliers and exceptional cases of institutions with atypical retention levels, such as Institution H, can address the most problematic cases concerning retention. Exploring different methodologies to estimate cut-off points is also an interesting research direction, particularly for its practical use to improve college retention.

Declaration of Competing Interest

All co-authors have seen and agree with the contents of the manuscript, and declare that they have no known competing financial interests, or personal relationships, that could have appeared to influence the work reported in this paper. We certify that the submission is original work and is not under review at any other publication.

Acknowledgements

This work was supported by the Consejo Nacional de Ciencia y Tecnología (CONACYT Grant # 32033), the National Science Foundation (NSF Grant # 1545667) and The Fulbright Commission of the US Department of State’s Fulbright Visiting Scholar Program.

References

Akera, A., Appelhans, S., Cheville, A., De Pree, T., Fatehiboroujeni, S., Karlin, J., & Riley, D. (2019). ABET & engineering accreditation - history, theory, practice: Initial findings from a national study on the governance of engineering education. *2019 ASEE Annual Conference & Exposition Proceedings*. https://doi.org/10.18260/1-2–32020, 32020.

Al Busaidi, H. (2020). Examining the relationship between the accreditation of engineering programmes and institutional performance. *Quality Assurance in Education, 28*(3), 179–192. https://doi.org/10.1108/QAE-10-2019-0098.

Aljohani, O. (2016a). A review of the contemporary international literature on student retention in higher education. *International Journal of Education and Literacy Studies, 4*(1), 40–52. https://doi.org/10.7575/aiac.ijels.v.4n.1p.40.

Aljohani, O. (2016b). A comprehensive review of the major studies and theoretical models of student retention in higher education. *Higher Education Studies, 6*(2), 1. https://doi.org/10.5539/hes.v6n2p1.

Astin, A. W. (1999). Student involvement: A developmental theory for higher education. *Journal of College Student Development, 40*(5), 12.

Barbera, S. A., Berkshire, S. D., Boronat, C. B., & Kennedy, M. H. (2020). Review of undergraduate student retention and graduation since 2010: Patterns, predictions, and recommendations for 2020. *Journal of College Student Retention: Research, Theory & Practice, 22*(2), 227–250. https://doi.org/10.1177/1521025117738233.

Becker, F. S. (2010). Why don’t young people want to become engineers? Rational reasons for disappointing decisions. *European Journal of Engineering Education, 35*(4), 349–366. https://doi.org/10.1080/03043797.2010.489941.

Belser, C. T., Shillingford, M. A., Daire, A. P., Prescod, D. J., & Dagley, M. A. (2018). Factors influencing undergraduate student retention in STEM majors: Career development, math ability, and demographics. *The Professional Counselor, 8*(3), 262–276. https://doi.org/10.15241/ctb.8.3.262.

Bennedsen, J., Adolfsen, N., Lauritsen, A. B., Edström, K., Kuptasthien, N., Roslöf, J., & Songer, R. (2019). *The 15th International CDIO Conference: Proceedings – Full Papers*. https://doi.org/10.7146/aul.347.

Berger, J. B., Blanco-Ramírez, G., & Lyons, S. (2012). Past to present: A historical look at retention. In A. Seidman (Ed.), *College student retention: Formula for student success* (pp. 7–34), 7–34.

Biancani, S., & McFarland, D. A. (2013). Social networks research in higher education. In M. B. Paulsen (Ed.), *Higher education: Handbook of theory and research* (Volume 28, pp. 151–215). Netherlands: Springer. https://doi.org/10.1007/978-94-007-5836-0_4.

Blat, C. E. (2018). *Mathematics readiness of entering college students*. The University of North Carolina at Charlotte.

Borrego, M., & Bernhard, J. (2011). The emergence of engineering education research as an internationally connected field of inquiry. *Journal of Engineering Education, 100*(1), 14–47. https://doi.org/10.1002/j.2168-9830.2011.tb00003.x.

Braxton, J. M., Doyle, W. R., III, H. V. H, Hirschy, A. S., Jones, W. A., & McLendon, M. K. (2013). *Rethinking college student retention*. John Wiley & Sons.

Braxton, J. M., Milem, J. F., & Sullivan, A. S. (2000). The influence of active learning on the college student departure process: Toward a revision of tinto's theory. *The Journal of Higher Education, 71*(5), 569–590. https://doi.org/10.1080/00221546.2000.11778853.

Burke, A. (2019). Student retention models in higher education: A literature review. *College and University, 94*(2), 12.

Cabrera, A. F., Castaneda, M. B., Nora, A., & Hengstler, D. (1992). The convergence between two theories of college persistence. *The Journal of Higher Education, 63*(2), 143. https://doi.org/10.2307/1982157.

Chapin, L. A. (2019). Longitudinal predictors for Mexican Americans' high school and college graduation: Individual and ecodevelopmental factors. *Journal of Latinos and Education*, 1–16. https://doi.org/10.1080/15348431.2019.1588733.

Cominole, M., Wheeless, S., Dudley, K., Franklin, J., & Wine, J. (2007). 2004/06 beginning postsecondary students longitudinal study (BPS:04/06). Methodology report. NCES 2008–2184. *National center for education statistics*. National Center for Education Statistics. https://eric.ed.gov/?q=coenrollment&ff1=subAcademic+Persistence&id=ED499336.

Committee on Prospering in the Global Economy of the 21st Century: An Agenda for American Science and Technology. (2007). *Rising above the gathering storm: Energizing and employing america for a brighter economic future* (p. 11463). National Academies Press. https://doi.org/10.17226/11463.

Corlu, M. S., Svidt, K., Gnaur, D., Lavi, R., Borat, O., & Corlu, M. A. (2018). Engineering education in higher education in Europe. In Y. J. Dori, Z. Mevarech, & D. Baker (Eds.), *Cognition, metacognition, and culture in STEM education* (pp. 89–116). Springer. https://doi.org/10.1007/978-3-319-66659-4_5.

Crisp, G. (2013). The influence of Co-enrollment on the success of traditional age community college students. *Teachers College Record, 115*(10).

Cunningham, A. F., & Milam, J. (2005). *Feasibility of a student unit record system within the integrated postsecondary education data system*. March 21. National Center for Education Statistics https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2005160.

Dinno, A. (2015). Nonparametric pairwise multiple comparisons in independent groups using Dunn's test. *The Stata Journal, 15*(1), 292–300. https://doi.org/10.1177/1536867X1501500117.

Dobryakova, M., & Froumin, I. (2010). Higher engineering education in Russia: Incentives for real change. *International Journal of Engineering Education, 26*, 1032–1041.

Dobson, A. J., & Barnett, A. G. (2008). *An introduction to generalized lienar models third edition* (3rd ed.). CRC Taylor & Francis Group.

Edström, K., & Kolmos, A. (2012). *Comparing two approaches for engineering education development: PBL and CDIO* (p. 16).

Edström, K., & Kolmos, A. (2014). PBL and CDIO: Complementary models for engineering education development. *European Journal of Engineering Education, 39*(5), 539–555. https://doi.org/10.1080/03043797.2014.895703.

ERIC. (2020). *Educational resources information center thesaurus of ERIC descriptors*. https://eric.ed.gov/.

Fluss, R., Faraggi, D., & Reiser, B. (2005). Estimation of the youden index and its associated cutoff point. *Biometrical Journal, 47*(4), 458–472. https://doi.org/10.1002/bimj.200410135.

Forsman, J., van den Bogaard, M., Linder, C., & Fraser, D. (2015). Considering student retention as a complex system: A possible way forward for enhancing student retention. *European Journal of Engineering Education, 40*(3), 235–255. https://doi.org/10.1080/03043797.2014.941340.

Forsman, J., Linder, C., Moll, R., Fraser, D., & Andersson, S. (2014). A new approach to modelling student retention through an application of complexity thinking. *Studies in Higher Education - STUD HIGH EDUC, 39*, 1–19. https://doi.org/10.1080/03075079.2011.643298.

Gardner, J. P., Brooks, C., & Li, W. (2018). Learn from your (Markov) neighbor: Coenrollment, assortativity, and grade prediction in undergraduate courses. *Journal of Learning Analytics, 5*(3). https://doi.org/10.18608/jla.2018.53.4.

Ginder, S. A. (2013). *2012-13 integrated postsecondary education data system (IPEDS) methodology report* (p. 24).

Grunspan, D. Z., Wiggins, B. L., & Goodreau, S. M. (2014). Understanding classrooms through social network analysis: A primer for social network analysis in education research. *CBE—Life Sciences Education, 13*(2), 167–178. https://doi.org/10.1187/cbe.13-08-0162.

Hodges, B., Bedford, T., Hartley, J., Klinger, C., Murray, N., O'Rourke, J., & Schofield, N. (2013). *Enabling retention: Processes and strategies for improving student retention in university-based enabling programs: Final report 2013*. Australian Government Office for Learning and Teaching.

Inzer, L. D., & Crawford, C. B. (2005). A review of formal and informal mentoring: Processes, problems, and design. *Journal of Leadership Education, 4*(1), 31–50.

Israel, U., Koester, B. P., & McKay, T. A. (2020). Campus connections: Student and course networks in higher education. *Innovative Higher Education, 45*(2), 135–151. https://doi.org/10.1007/s10755-019-09497-3.

Jesiek, B. K., Newswander, L. K., & Borrego, M. (2009). Engineering education research: Discipline, community, or field? *Journal of Engineering Education, 98*(1), 39–52. https://doi.org/10.1002/j.2168-9830.2009.tb01004.x.

Johnson, M. H. (2013). *An analysis of retention factors in undergraduate degree programs in science, technology, engineering, and mathematics*.

Johnson, W. C., & Jones, R. C. (2006). Declining interest in engineering studies at the time of increased business need. In L. E. Weber, & J. J. Duderstadt (Eds.), *Universities and business: Partnering for the knowledge society* (pp. 243–252). London: Economica.

Jones, D. R. (2008). *Student retention and success: A synthesis of research* (p. 43).

Kang, Z. (2019). Using machine learning algorithms to predict first-generation college students' six-year graduation: A case study. *International Journal of Information Technology and Computer Science, 11*(9), 1–8. https://doi.org/10.5815/ijitcs.2019.09.01.

Kayis, B. (2004). Trends in undergraduate mechanical, manufacturing, aerospace and mechatronics programs worldwide: In-depth study of 55 universities' programs. In *2004 IEEE International Engineering Management Conference (IEEE Cat. No.04CH37574), 1* pp. 178–182). https://doi.org/10.1109/IEMC.2004.1407100.

Krause, K.-L., & Armitage, L. (2014). Australian student engagement, belonging, retention and success: A synthesis of the literature. *The Higher Education Academy*, 1–45.

Lee, Y.-G., & Ferrare, J. J. (2019). Finding one's place or losing the race? The consequences of STEM departure for college dropout and degree completion. *The Review of Higher Education, 43*(1), 221–261. https://doi.org/10.1353/rhe.2019.0095.

Lin, C.-H. (2020). Making graduation matter: Initial evidence and implications for policy and research. *Studies in Higher Education*, 18.

Lord, S. M., Ohland, M. W., Layton, R. A., & Camacho, M. M. (2019). Beyond pipeline and pathways: Ecosystem metrics. *Journal of Engineering Education, 108*(1), 32–56.

Lucena, J., Downey, G., Jesiek, B., & Elber, S. (2008). Competencies beyond countries: The Re-organization of engineering education in the United States, Europe, and Latin America. *Journal of Engineering Education, 97*(4), 433–447. https://doi.org/10.1002/j.2168-9830.2008.tb00991.x.

Malmqvist, J., Hugo, R., & Kjellberg, M. (2015). *A survey of CDIO implementation globally – Effects on educational quality* (p. 17).

Melguizo, T. (2011). A review of the theories developed to describe the process of college persistence and attainment. In J. C. Smart, & M. B. Paulsen (Eds.), *Higher education: Handbook of theory and research* (Vol. 26, pp. 395–424). Netherlands: Springer. https://doi.org/10.1007/978-94-007-0702-3_10.

NCES, N. R. (2019). *Persistence & Retention—2019*. July 10. National Student Clearinghouse Research Center https://nscresearchcenter.org/snapshotreport35-first-year-persistence-and-retention/.

Nicoletti, M. do C. (2019). Revisiting the Tinto's theoretical dropout model. *Higher Education Studies, 9*(3), 52. https://doi.org/10.5539/hes.v9n3p52.

Niu, L. (2020). A review of the application of logistic regression in educational research: Common issues, implications, and suggestions. *Educational Review, 72*(1), 41–67. https://doi.org/10.1080/00131911.2018.1483892.

Ohland, M. W., & Long, R. A. (2016). *The multiple-institution database for investigating engineering longitudinal development: An experiential case study of data sharing and reuse* (p. 25).

Ohland, M. W., Brawner, C. E., Camacho, M. M., Layton, R. A., Long, R. A., Lord, S. M., & Wasburn, M. H. (2011). Race, gender, and measures of success in engineering education. *Journal of Engineering Education, 100*(2), 225–252.

Ohland, M. W., Sheppard, S. D., Lichtenstein, G., Eris, O., Chachra, D., & Layton, R. A. (2008). Persistence, engagement, and migration in engineering programs. *Journal of Engineering Education, 97*(3), 259–278.

Orr, M. K., Ohland, M. W., Lord, S. M., & Layton, R. A. (2020). Comparing the multiple-institution database for investigating engineering longitudinal development with a national dataset from the United States. *International Journal of Continuing Engineering Education, 36*(4), 1321–1332.

Passow, H. J., & Passow, C. H. (2017). What competencies should undergraduate engineering programs emphasize? A systematic review. *Journal of Engineering Education, 106*(3), 475–526. https://doi.org/10.1002/jee.20171.

Rao, A. R., & Bandyopadhyay, S. (1987). Measures of reciprocity in a social network. *Sankhyā: The Indian Journal of Statistics, Series A*, 141–188.

Rintala, U., Andersson, S., & Kairamo, A.-K. (2021). *How to retain students in higher engineering education? Findings of the ATTRACT project*.

Rintala, U., Kairamo, A.-K., Andersson, S., Strandås, M., Kelly, K., Gonçalves, I., & Lucas, A. (2012). *Knowing our students – Different approaches to student retention* (p. 8).

Seidman, A. (Ed.). (2005). *College student retention: Formula for student success*. Greenwood Publishing Group.

Sithole, A., Chiyaka, E. T., McCarthy, P., Mupinga, D. M., Bucklein, B. K., & Kibirige, J. (2017). Student attraction, persistence and retention in STEM programs: Successes and continuing challenges. *Higher Education Studies, 7*(1), 46. https://doi.org/10.5539/hes.v7n1p46.

Thomas, S. L. (2000). Ties that bind: A social network approach to understanding student integration and persistence. *The Journal of Higher Education, 71*(5), 591. https://doi.org/10.2307/2649261.

Tight, M. (2020). Student retention and engagement in higher education. *Journal of Further and Higher Education, 44*(5), 689–704. https://doi.org/10.1080/0309877X.2019.1576860.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*(1), 37.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (second edition, p. 60637). 5801 South Ellis Avenue, Chicago, IL: University of Chicago Press.

Tinto, V. (1997). Classrooms as communities: Exploring the educational character of student persistence. *The Journal of Higher Education, 68*(6), 599. https://doi.org/10.2307/2959965.

Tinto, V. (2006). *Research and practice of student retention: What next?* (p. 19). https://doi.org/10.2190/4YNU-4TMB-22DJ-AN4W.

Titus, M. A. (2004). An examination of the influence of institutional context on student persistence at 4-year colleges and universities: A multilevel approach. *Research in Higher Education, 45*(7), 673–699. https://doi.org/10.1023/B:RIHE.0000044227.17161.fa.

Wang, X., & McCready, B. (2013). The effect of postsecondary coenrollment on college success: Initial evidence and implications for policy and future research. *Educational Researcher, 42*(7), 392–402. https://doi.org/10.3102/0013189X13505683.

Wang, X., & Wickersham, K. (2014). Postsecondary co-enrollment and baccalaureate completion: A look at both beginning 4-year college students and baccalaureate aspirants beginning at community colleges. *Research in Higher Education, 55*(2), 166–195. https://doi.org/10.1007/s11162-013-9317-4.

Willcoxson, L., Cotter, J., & Joy, S. (2011). Beyond the first-year experience: The impact on attrition of student experiences throughout undergraduate degree studies in six diverse universities. *Studies in Higher Education, 36*(3), 331–352.

Zhang, G., Anderson, T. J., Ohland, M. W., & Thorndyke, B. R. (2004). Identifying factors influencing engineering student graduation: A longitudinal and cross-institutional study. *Journal of Engineering Education, 93*(4), 313–320. https://doi.org/10.1002/j.2168-9830.2004.tb00820.x.