

Does Curricular Complexity Imply Program Quality?

Prof. Greg L. Heileman, University of Kentucky

Gregory L. Heileman received the BA degree from Wake Forest University in 1982, the MS degree in Biomedical Engineering and Mathematics from the University of North Carolina-Chapel Hill in 1986, and the PhD degree in Computer Engineering from the University of Central Florida in 1989. In 1990 he joined the Department of Electrical and Computer Engineering at the University of New Mexico, Albuquerque, NM, where he is currently a Professor. Since 2011 he has served as the Associate Provost for Curriculum at the University of New Mexico. During 1998 he held a research fellowship at the Universidad Carlos III de Madrid, and in 2005 he held a similar position at the Universidad Politénica de Madrid. His research interests are in information security, the theory of computing and information, machine learning, and data structures and algorithmic analysis. He is the author of the text Data Structures, Algorithms and Object-Oriented Programming, published by McGraw-Hill in 1996.

William G. Thompson-Arjona, University of Kentucky

Will Thompson is a graduate student in the department of electrical and computer engineering at the University of Kentucky. Prior to this, he was a hardware development engineer in the industrial automation sector working for Rockwell Automation (NYSE: ROK). He earned a B.S. in bioelectrical engineering from Marquette University in 2015. His interests include optimization, embedded hardware systems, signal processing, and machine learning.

Mr. Orhan Abar, University of Kentucky

Orhan Abar is a Ph.D. student in the Computer Science (CS) Department at the University of Kentucky. He graduated with an M.S in CS from the University of Texas at San Antonio in 2013. He obtained a B.S. in Computer Engineering from Firat University in 2009. His research interests include Deep Learning, Data Mining, Frequent Pattern Mining, and Optimization.

Mr. Hayden W. Free, University of Kentucky

Hayden Free is an undergraduate student studying Computer Science at the University of Kentucky. His focused area of interests include distributed systems, cloud architecture, and software design.

Does Curricular Complexity Imply Program Quality?

Gregory L. Heileman[†], William G. Thompson-Arjona[†], Orhan Abar[‡] and Hayden W. Free[‡]

{greg.heileman, wgthompson, orhan.abar, hayden.free}@uky.edu

[†]Department of Electrical & Computer Engineering

[‡]Department of Computer Science

University of Kentucky

Abstract

A number of metrics exist for quantifying the complexity of academic program curricula. Complexity in this case relates the extent to which the structure of a curriculum impacts a student's ability to progress through that curriculum towards graduation. The ability to quantify curricular complexity in this manner allows us to order programs according to their complexity, and to compare and contrast similar programs at different institutions according to these complexity measures. When sharing this type of information with faculty and program administrators, those at programs at the higher end of the complexity scale often speculate that high complexity implies a higher quality program. Which leads to the more general question, what does curricular complexity tell us about program quality? In cursory investigations of this conjecture, a surprising relationship emerged. Specifically, anecdotal review provided significant evidence to support the proposition that higher quality engineering programs have lower complexity curricula. It is worth noting that if this proposition is indeed true, then the contrapositive proposition, that higher complexity curricula imply lower quality programs also holds. In this study we collected a sufficient amount of data to determine the veracity of this proposition for undergraduate electrical engineering programs.

The methodology employed in this study involved partitioning a large set of undergraduate electrical engineering curricula into three categories (top tier, mid tier and bottom tier) according to their quality. The curricular complexity variance within and between these groups was then analyzed using ANOVA methodologies. Because program quality is a subjective measure, we used the 2018 U.S. News & World Report Best Undergraduate Engineering Program rankings as a proxy for quality. The first group included schools in the top decile of this ranking, the medium group included schools from the fourth and fifth deciles, and the low group included those schools that were grouped together at the bottom of the list (approximately the bottom decile). The null hypothesis was that there are no significant differences between the intragroup and intergroup curricular complexity measures. This analysis found that with a low margin of error, and a 95% confidence interval, the null hypothesis should be rejected. Furthermore, the most significant difference was between the set of highly-ranked programs and the lowest-ranked programs, with a less pronounced difference between the medium- and lowest-ranked programs. It is generally the case that higher ranked schools admit better prepared students, and they have more resources available to support these students than do lower ranked schools. Thus, we expect the students at higher ranked schools to graduate at higher rates than those at lower ranked schools. The study reported in this paper shows that electrical engineering undergraduate students at higher ranked schools receive another student success advantage; namely, they encounter less complex curricula.

Introduction

Over the past few years we have developed a number of metrics for quantifying the complexity of academic program curricula.^{5,7,11} This has led to the development of a *curricular complexity metric* that directly relates how the structure of a curriculum impacts a student's ability to progress through that curriculum to graduation. We refer to the general study of how program curricula impact student academic success as *curricular analytics*. Research in this area demonstrates that according to these complexity metrics, engineering programs tend to be among the most complex curricula at a university. This is attributed to the large number of prerequisites that accompany many of the courses in engineering programs, as well as the long prerequisite chains that tend to exist in these curricula.

To gain a better understanding of the aforementioned factors, consider the electrical engineering degree plan shown in Figure 1, offered by a university in the southwest of the United States that has a high curricular complexity score. The analysis provided in this figure was created by utilizing the Curricular Analytics Toolbox, an open source framework created for the purpose of analyzing university curricula.⁶ The complexity associated with a given course c is a function of the number of courses that are "blocked" by c (i.e., the number of courses that cannot be attempted until c is successfully completed), and the longest path in the curriculum that c is on. The complexity associated with a term is given by the sum of the complexities of the courses in the term, and the complexity of a curriculum is given by the sum of the complexities of the courses in the curriculum. We previously demonstrated a direct relationship between the complexity of a curriculum and a student's ability to complete that curriculum.⁵

Curricular analytics metrics have been used by curriculum committees to assess the potential impact of particular curricular reforms, and to compare similar programs at different schools. These analyses reveal that in many disciplines there exist significant variances in curricular complexity between the curricula implemented at different institutions. For instance, in Figure 2 we show the electrical and computer engineering degree plan at a university located in the northeast of the United States, a program with a low curricular complexity score.

What is remarkable about the two programs shown in Figures 1 and 2 is that they have nearly identical student learning outcomes. In particular, both programs list the eleven ABET student learning outcomes, and they both include the additional program requirements ABET stipulates should accompany electrical, computer, communications, telecommunication(s) and similarly named programs.⁴ Furthermore, both programs have ABET accreditation. Thus, experts in the discipline have independently certified the quality of these programs and have determined that both programs produce graduates who successfully attain the required learning outcomes. However, from a structural



Figure 1: The eight-term degree plan for the electrical engineering bachelor of science program at a university located in the southwest for the 2017-18 academic year. The eighth-term Senior Design course (ECE 4336) is selected in this degree plan. The highlighted courses, 25 in total, must all be successfully completed before a student may attempt ECE 4336. The longest path in this curriculum, shown as a dashed blue line, starts at Calculus I (MATH 1431), terminates at ECE 4336, and includes twelve total courses. The complexity of a course is given by the number inside the course vertex and the complexity of a term is shown beneath each term. The complexity of the entire curriculum is 487. This visualization was created using the Curricular Analytics Toolbox.⁶



Figure 2: The eight-term degree plan for the electrical and computer engineering bachelor of science program at a university located in the northeast for the 2017-18 academic year. The fifth-term Intelligent Physical Systems course (ECE 3400) is selected in this degree plan. In this curriculum, there are ten courses that must be successfully completed before a student may attempt ECE 3400, and ECE 3400 is on one of the longest paths in the curriculum, which includes five courses. The complexity of a course is given by the number inside the course vertex and the complexity of a term is shown beneath each term. The complexity of the entire curriculum is 118. This visualization was created using the Curricular Analytics Toolbox.⁶

perspective, it is readily apparent that these two curricula are significantly different. Indeed, the structural arrangement of the courses and the prerequisites produce a curricular complexity metric for the electrical engineering program shown in Figure 1 that is more than four times the curricular complexity metric of the electrical and computer engineering program shown in Figure 2.

When sharing curricular analyses of the type shown in Figures 1 and 2 with engineering program faculty and administrators, some at programs with higher curricular complexity have speculated that higher complexity may be an indicator of higher program quality. The experiments described in the following sections were designed to address this speculation. More specifically, in the following sections we describe a set of experiments that addresses this question for the case of undergraduate electrical engineering programs at doctoral-granting institutions.

An interesting feature of this study is that nearly every program considered has the same ABET accreditation. Thus, as mentioned above, they share the same set of student learning outcomes, and are therefore equivalent in the sense that they all produce competent engineers. Furthermore, the courses that are offered appear very similar across these curricula. For instance, they all appear to include Calculus I in the first term of the freshman year, and a Circuits course in the sophomore year, etc. There is, however, extreme variability in the way different electrical engineering programs structure their curricula. Some programs include a larger number of courses (and credit hours) and are tightly prescribed in that they stipulate a larger number of prerequisites for key courses, as exemplified by the curriculum in Figure 1. Other programs have fewer courses (e.g., they meet the 120-credit-hour minimum that regional accreditors expect) and they provide more freedom by having fewer prerequisites for key courses, and a proportionally larger number of elective courses, as exemplified by the curriculum in Figure 2.

Methodology

A formal and rigorous assessment of the "quality" of engineering programs that could be used as a basis for fine-grained comparisons of programs requires one to establish a set of quality metrics, as well as a means for fairly and accurately assessing them. Even arriving at agreement on a set of quality metrics is problematic, as most programs would argue for the creation of metrics (and a weighting of these metrics) that align with their particular institutional values and emphases. Thus, no agreed upon nationally or internationally normed quality rubric that can be used to compare engineering programs at a sufficient level of detail is likely to be developed. ABET accreditation standards, on the other hand, have been created to ensure that engineering programs operate above a certain quality threshold. An engineering program either receives ABET accreditation or not, there are no "degrees" of ABET accreditation that might be used to constitute a ranking. An opinionated ranking of engineering programs according to their quality is therefore the best that we can hope for.

The most well-known rankings in higher education are conducted by the U.S. News & World Report. In this study we used the rankings provided by the U.S. News & World Report 2018 Best Undergraduate Engineering Programs survey as a proxy for program quality. That is, for the purpose of this study, we assume that the "best" engineering programs are synonymous with the highest quality engineering programs. We acknowledge the concerns routinely expressed regard-

ing these rankings.^{1,2} However, it should be noted that this study uses aggregations of schools within tiers, and the statistics associated with these aggregations. Thus, the specific rankings of the schools within the tiers are irrelevant, all that matters is the tier in which a school is placed. Upon inspection of the schools within each tier, we believe that knowledgable and impartial observers would agree that the three tiers constructed in this study are highly correlated with program quality.

To appear on a U.S. News & World Report undergraduate engineering survey, a school must have at least one undergraduate engineering program that is accredited by ABET. Two surveys are conducted, one for schools whose highest engineering degree offered is a doctorate and another for schools whose highest engineering degree offered is a bachelor's or master's. For this study we consider the ranking provided by the former; that is, the ranking of doctoral institutions. This ranking is based solely on the peer assessment provided by deans and senior faculty members at doctoral institutions, and involves asking these survey participants to rate each program they are familiar with on a scale from 1 (marginal) to 5 (distinguished). Two peer assessment surveys are sent to each ABET-accredited engineering program at these schools, with a response rate of approximately 58%. U.S. News & World Report uses the two most recent years' responses to calculate weighted average scores of programs, which determines the ranking. The U.S. News & World Report 2018 Best Undergraduate Engineering Programs rankings for doctoral institutions lists 205 schools, with a formal ranking designation given to the programs in the 1-177 range. Programs ranked lower than 177 are listed alphabetically and lumped into a ranking category.

Experiment Design

The question of interest in this study is whether or not curricular complexity is related to program quality. In order to answer this question, we constructed an analysis of variance (ANOVA) experiment that involved partitioning the schools in the U.S. News & World Report 2018 Best Undergraduate Engineering Programs rankings according to their decile within the ranking. From these deciles, three tiers were created as follows. A *top tier* of schools defined as those in the first decile of the ranking. A *mid tier*, defined as the set of schools in the fifth and sixth deciles of the ranking that are equidistant from first to last ranked schools. A *bottom tier* of schools comprised of the schools ranked below 177, which spans a little more than one decile at the bottom of the ranking. The null hypothesis is, "there is no difference between the mean values of the curricular complexities of those schools belonging to the top, mid and bottom tiers." The alternative hypothesis is then, "at least one of the curricular complexity mean values of a school tier differs significantly from the means of the other two tiers."

The ANOVA analysis involves random sampling of schools within each of the three tiers. In order to ensure the analysis is able to distinguish between actual curricular complexity differences among the tiers, and random variation, sufficient sample sizes must be determined. Under the assumption that the curricular complexity distributions within the tiers are approximately normal, with variance σ^2 , the number of samples that should be selected from each tier is given by

$$n = \left(\frac{\sigma Z}{E}\right)^2,\tag{1}$$

where Z is the confidence interval expressed using deviation within the standard normal distribution, and E is the margin of error. To obtain an estimate of σ , pilot samples from each of the three tiers were taken, yielding the estimate $\hat{\sigma} = 90$. For a 95% confidence interval, which corresponds to Z = 1.96, the margin of error will be 40 curricular complexity points, i.e., 20 points on either side of the mean for a tier. Using these values in Equation (1) leads to sample sizes of $n_1 = n_2 = n_3 = 20$, where n_1, n_2 and n_3 are the sample sizes for the top, medium and bottom tiers, respectively. Thus, by sampling at least 20 schools from each tier, we can have 95% confidence that the error in this analysis will be by no more than 40 curricular complexity points.

In order to test the null hypothesis using ANOVA, we must assume the curricular complexity values of the schools within each tier are normally distributed, and that all three tiers have the same variance σ^2 . It should be noted that these conditions can be moderately relaxed (particularly the normality assumption) and the analysis will remain valid.⁹

The ANOVA method partitions the total sum of squares of the deviations in curricular complexity across all schools into two independent parts, one that is attributed to the independent variable (program quality in this case), and a remainder that is attributed to random errors arising from other factors not accounted for in this experiment. That is,

$$TSS = SST + SSE,\tag{2}$$

where TSS denotes the total sum of squares of deviations, SST represents the sum of squares of the deviations between the tiers, and SSE is the sum of squares attributed to errors or noise. More specifically, if we let cc_{ij} denote the curricular complexity of the j^{th} school sampled from the i^{th} tier, then

$$TSS = \sum_{i=1}^{3} \sum_{j=1}^{n_i} \left(cc_{ij} - \overline{cc} \right)^2, \tag{3}$$

where \overline{cc} is the sample mean for all samples drawn over all tiers. The sum of squares deviation between the tiers is given by

$$SST = \sum_{i=1}^{3} n_i \left(\overline{T}_i - \overline{cc}\right)^2,\tag{4}$$

where T_i is the total curricular complexity of the schools sampled from the i^{th} tier, and $\overline{T}_i = T_i/n_i$, i = 1, 2, 3, are the tier sample averages. Note that when the sample means for the three tiers are the same, SST = 0.

Substituting Equations (3) and (4) into Equation (2) and solving for SSE yields:

$$SSE = \sum_{i=1}^{3} \sum_{j=1}^{n_i} \left(cc_{ij} - \bar{T}_i \right)^2.$$
(5)

The unbiased estimator of σ^2 based on n-3 degrees of freedom is given by the mean square error,

$$MSE = \frac{SSE}{n-3},\tag{6}$$

where $n = n_1 + n_2 + n_3$. The mean square for the tiers has 2 degrees of freedom, i.e., one less than the number of tiers, and is therefore

$$MST = \frac{SST}{2}.$$
 (7)

In order to assess the statistical significance of a decision to reject the null hypothesis, an *F*-test is conducted to compare the deviation among the tier variances. The *F*-test statistic is given by

$$F = \frac{MST}{MSE}.$$

Note that the *F*-test is a ratio that compares the mean square variability between the tiers to the mean square variability within the tiers. Thus, as *F*-test values increase above 1, the data are increasingly inconsistent with the null hypothesis, and the null hypothesis should be rejected when $F > F_{\alpha}$, where F_{α} is the critical value of *F* where the probability of a type I error is α .

For the *F* distribution with (2, 65) degrees of freedom, $F_{0.05} = 3.15$. That is, if the *F*-test for the experiment yields a value greater than 3.15, we can reject the null hypothesis with only a 5% chance of doing so in error.

Results

According to the sample size analysis provided above, at least 20 schools were randomly sampled from each tier. From the top decile of the News & World Report ranking, 21 schools were sampled, which constitute nearly the entirety of the schools in this tier, and include: California Institute of Technology, Carnegie Mellon University, Columbia University, Cornell University, Duke University, Georgia Institute of Technology, Johns Hopkins University, Northwestern University, Princeton University, Purdue University, Rice University, Stanford University, Texas A&M University–College Station, University of California–Berkeley, University of California–Los Angeles, University of California–San Diego, University of Illinois–Urbana-Champaign, University of Michigan–Ann Arbor, University of Texas–Austin, University of Wisconsin and Virginia Polytechnic Institute and State University.

A total of 21 schools were sampled from the middle tier of the U.S. News & World Report ranking, including: Brigham Young University, Clarkson University, Embry-Riddle Aeronautical University, George Washington University, Indiana University-Purdue University–Indianapolis, Louisiana State University–Baton Rouge, New Jersey Institute of Technology, Oklahoma State University, San Diego State University, Southern Methodist University, Texas Tech University, University of Alabama, University of California–Riverside, University of California–Santa Cruz, University of Cincinnati, University of Houston, University of Kentucky, University of Miami, University of Missouri, University of North Carolina–Charlotte and University of Oklahoma.

Finally, 21 schools were sampled from the bottom tier of the News & World Report ranking. These schools include: Florida Atlantic University, Jackson State University, Lamar University, Morgan State University, Oakland University, Prairie View A&M University, South Dakota State



Figure 3: The curricular complexity histogram for all schools included in the study. The average complexity value of these schools is 273.6, with a standard deviation of 104.2.



Figure 4: The curricular complexity histograms of the schools in the study, disaggregated by tier. Note that these are approximately normal with similar variances. The top tier sample has an average curricular complexity of 188 with a standard deviation of 90. The mid tier sample has an average curricular complexity of 285 with a standard deviation of 81. The bottom tier sample has an average curricular complexity of 337 with a standard deviation of 76.

University, Tennessee State University, Texas A&M University–Kingsville, Texas State University, Tuskegee University, University of Bridgeport, University of Denver, University of Detroit Mercy, University of Louisiana–Lafayette, University of Missouri–Kansas City, University of New Orleans, University of North Dakota, University of North Texas. University of Tennessee–Chattanooga, Western Michigan University.

A histogram showing the curricular complexity distribution of all schools included in the study (i.e., the schools sampled to form all three tiers) is provided in Figure 3. The distribution of school complexities in this figure appears approximately Gaussian with $\mu = 273.6$ and $\sigma = 104.2$. However, when the school complexities are disaggregated by the previously defined tiers, as shown in Figure 4, possible differences appear.

Box-and-whisker diagrams for each of the three samples, provided in Figure 5, shows the differences between the various statistics associated with these samples. These diagrams again indicate that there are possible curricular complexity differences between the tiers. The question is whether



Figure 5: Box-and-whisker diagrams for each of the three samples taken from each of the three defined tiers. The box for each tier encompasses the upper (75%) and lower (25%) quartiles, i.e., the interquartile range, of the curricular complexities of the schools in the sample, the line inside the box is the median value of the sample, and the whiskers show the extreme curricular complexity scores (excluding outliers) within each sample. Outliers (a curricular complexity score greater/less than 1.5 times the upper/lower quartile) are shown as dots in the diagrams.

	Sum of Squares	Deg. of Freedom	Mean Square	F
Tiers	238494	2	119247	17.45
Error	409932	60	6832	
Total	648426	62		

Table 1: The results of the ANOVA analysis associated with the samples selected from the three tiers of schools. The F-test statistic is 16.38.

or not the curricular complexity differences between the tiers are statistically significant. The ANOVA analysis described in the prior section was applied to answer this question. Table 1 shows the ANOVA statistics that resulted from applying the sampled data, including the outliers. Notice that the F-test statistic obtained from this analysis is 17.45. Because

$$17.45 > F_{0.05} = 3.15,$$

the null hypothesis should be rejected. That is, with a low probability of error, the samples collected from each tier indicate that the mean curricular complexity values of the tiers are different. This result, along with the evidence given in Figures 4 and 5, provide strong evidence that higher quality electrical engineering programs have lower curricular complexity, and that lower quality electrical engineering programs have higher curricular complexity.

Discussion

We have demonstrated that an inverse relationship exists between the complexity of the curricula in undergraduate electrical engineering programs and the perceived quality of these programs.

Specifically, at doctoral-granting engineering schools the complexity of the electrical engineering undergraduate curricula at the highest quality schools (where quality is subjectively determined by a survey proved to all schools in this category) is drastically less than the complexity of the curricula at those schools judged to be at the lower end of this quality ranking. The average complexity of those schools at the bottom of the raking is almost twice the average of those schools in the top decile of the ranking. In addition, we demonstrated that this difference is statistically significant; that is, this difference is due to something other than chance. Because the complexity of a curriculum is a measure of the difficulty that students are expected to have completing that curriculum, this difference has important student success implications. In particular, if we were to equalize instructional factors (e.g., the difficulty of the courses in the curriculum, the support services provided to students, etc.) and student background preparation, we would expect students to graduate at a higher rate from the lower complexity curricula. This is indeed what we observe, and the fact that this benefit is most pronounced in the highest quality programs deserves further investigation. It is possible to produce arguments for both sides of a possible cause-and-effect relationship between curricular complexity and program quality. Below we consider a few.

One might argue that because the top tier schools admit better prepared students they can offer less complex curricula, as their students can more easily overcome any knowledge gaps that may exist due to having fewer prerequisites prior to attempting a given course, as well as fewer total courses in their curricula. It should be noted, however, that there are a number of schools outside the tier that have created pathways in the first year of their curricula that substantially reduce curricular complexity.⁷ These curricular innovations have been demonstrated to significantly improve graduation rates, as well as the attainment of program learning outcomes.⁸ That is, it is possible to reduce the complexity of engineering programs that serve less-prepared students, while actually improving program quality (as judged by outcomes). More generally, we note that the principle of Occam's razor is often applied to guide engineering designs towards the simplest and therefore best solutions. One of the most popular versions of this principle states, "Entities are not to be multiplied without necessity." We posit that this study indicates this principle applies to curricula. Namely, the simplest curriculum (in terms of complexity) that allows students to attain a program's learning outcomes yields the best student success outcomes and therefore the highest quality program.

References

¹ M. N. Bastedo and N. A. Bowman. U.S. News & World Report college rankings: Modeling institutional effects on organizational reputation. *American Journal of Education*, 116:163–183, 2010.

² J. Byrne. Forbes magazine. How Much Attention Should You Pay To U.S. News' College Rankings?, Sept. 10, 2018.

³Electrical and Computer Engineering, Program Requirements. Cornell University. www.ece.cornell.edu/ece/programs/undergraduate-programs/majors/program-requirements, 2017.

- ⁴ Engineering Accreditation Commission, ABET. Criteria for accrediting engineering programs: Effective for reviews during the 2017-2018 accreditation cycle. www.abet.org, October 2016.
- ⁵ G. L. Heileman, C. T. Abdallah, A. Slim, and M. Hickman. Curricular analytics: A framework for quantifying the impact of curricular reforms and pedagogical innovations. *www.arXiv.org*, arXiv:1811.09676 [cs.CY], 2018.
- ⁶G. L. Heileman, H. W. Free, O. Abar, and W. G. Thompson-Arjona. Curricular Analytics Toolbox. https://github.com/CurricularAnalytics/CurricularAnalytics.jl.
- ⁷ G. L. Heileman, M. Hickman, A. Slim, and C. T. Abdallah. Characterizing the complexity of curricular patterns in engineering programs. In *Proceedings of the 2017 American Society for Engineering Education (ASEE) Annual Conference*, 2017.
- ⁸ N. W. Klingbeil and A. Bourne. The Wright State model for engineering mathematics education: Longitudinal impact on initially underprepared students. In *Proceedings of the 122nd ASEE Annual Conference*, Seattle, WA, June 14–17, 2015.
- ⁹ W. Mendenhall, R. L. Scheaffer, and D. D. Wackerly. *Mathematical Statistics with Applications*. Duxbury Press, Boston, MA, 2nd edition, 1981.
- ¹⁰ University of Houston. 2017-18 Undergradaute Catalog. http://publications.uh.edu/index.php?catoid=25, 2017.
- ¹¹ J. Wigdahl, G. L. Heileman, A. Slim, and C. T. Abdallah. Curricular efficiency: What role does it play in student success? In 2014 ASEE Annual Conference & Exposition, Indianapolis, Indiana, June 2014. ASEE Conferences. https://peer.asee.org/20235.