CRITICAL THINKING

SPRING, SUMMER AND FALL 2020

Report prepared by Ana Kriletic, MS (Graduate Assistant for Data Analytics) and Katie Boyd, PhD (Director, Academic Assessment)

EXECUTIVE SUMMARY

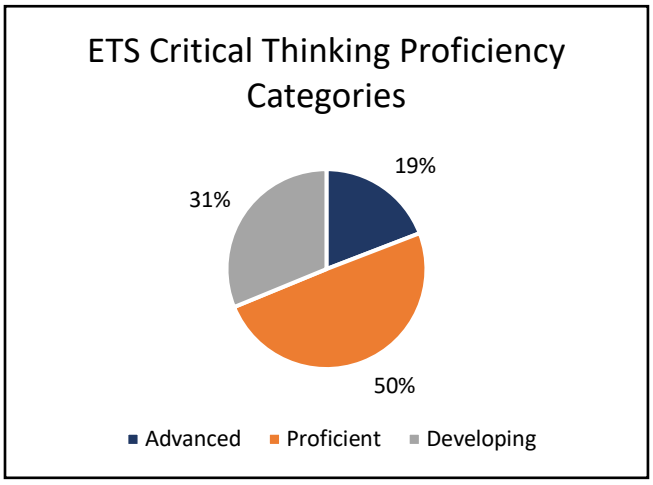
Sample

In total, 5151 students took the ETS HEIghten Critical Thinking Assessment as part of the SCORE (Student Core Outcomes and Readiness Evaluation) in 2020. The SCORE is embedded in a required senior graduation course, Creed to Succeed. Students participate in all requirements at very high rates (>90%) therefore, the sample is demographically very similar to all seniors that graduated in 2020. Unmotivated students were identified based on self-reported effort and time to complete test estimates. Additionally, students that received lowest score possible on all three Critical Thinking scores were removed prior to analysis. In total, 823 students (15.98%) were identified as 'unmotivated' and removed from the sample, resulting in 4328 usable cases.

Overall Performance

Figure 1 displays that overall, 69% of students were classified as "Proficient" or "Advanced" based on their performance on the assessment; our sample generally did better than the normed ETS sample group (although the normed group included data from freshmen, sophomores, juniors, and seniors). Subsequent exploratory analyses revealed statistically significant gaps in knowledge, described below.

Figure 1 – Overall Critical Thinking Scores



Gaps in Knowledge

Our hope is to create improvement initiatives where gaps in student learning are identified. For critical thinking, gaps were only identified among transfer students. Specifically:

Transfer: Students who began their higher education career at Auburn as freshmen outperformed transfer students on all aspects of the Critical Thinking test. However, further analysis showed that these differences were due to differences in ACT scores.

Low ACT scores: ACT scores were the best predictor of performance on Critical Thinking assessment such that students with higher ACT scores outperformed students with lower ACT scores. Thus, potential critical thinking interventions should target all students with low ACT scores and not just transfer students.

SCORE

TABLE OF CONTENTS

BACKGROUND

CHANGES TO GENERAL EDUCATION ASSESSMENT

From 2011 until 2015, general education was assessed through a "course-embedded" assessment approach. That is, faculty teaching core courses were asked to evaluate student work in their courses using a rubric developed by the Core Curriculum General Education Committee (CCGEC). There were eleven student learning outcomes and associated rubrics. In Fall 2015, the CCGEC began a year of reflection in which they met with faculty across campus to explore the effectiveness of the course-embedded assessment approach. Generally, the committee found that this approach was not working well for formative or summative assessment purposes. In 2016, the CCGEC began exploring other assessment options in an effort to centralize assessment and focus on graduating seniors. With this new focus, faculty working groups around each student learning outcome were tasked with (1) re-stating the student learning outcome with graduating seniors in mind, and (2) identifying, developing, or refining a measure aligned with the student learning outcome. All student learning outcomes were finalized by the CCGEC and approved by the University Senate in October 2017.

OUTCOME, ASSESSMENT, AND ALIGNMENT

In 2011 there were two outcomes related to critical thinking. Specifically, the original version of these outcomes written in 2011 stated: (2) Students will be able to read analytically and critically; and (3) Students will be able to critique and construct an argument effectively. The working group for this outcome was composed of faculty from the Departments of English (Dr. Miriam Clark) and Philosophy (Dr. James Shelley and Dr. Michael Watkins). To meet their charge, the working group met bi-weekly during the Fall 2016 semester and again during the Spring 2017 semester and drafted a new, single outcome statement. Specifically, outcomes 2 and 3 were refined to: "Students will be able to read and think critically." The CCGEC, to create consistency, developed a preface statement for the set of outcomes; thus, the current outcome reads in totality: ***"In order to become lifelong learners and use their education to solve practical problems, by the time of graduation, students will be able to effectively… read and think critically."***

The working group also determined 3 sub-outcomes:

Students will be able to:

1. identify the genre of the text, make reasonable inferences about its central purpose or argument, define its key components, and show how the writer uses these to reach a conclusion or create meaning or impact.
2. engage the text dialogically, questioning its premises, identifying its limitations, or advancing alternative perspective.
3. construct a strong, well-reasoned argument by determining which conclusion is supported by the strongest evidence.

To evaluate this outcome and the sub-outcomes, the working group decided to use the HEIghten Critical Thinking test, developed by ETS©. The HEIghten produces the following scores: Analytic, Synthetic, and an Overall score. Our Sub-outcome 1 is aligned with the HEIgthen tests' "Analytic" score. Sub-outcome 2 is aligned with the Overall Critical Thinking score, and the third sub-outcome is aligned with the "Synthetic" score.

SCORE TESTING & COMPLIANCE

All graduating seniors have an 'AT Hold' placed on their account when they enroll in UNIV 4AA0 – Creed to Succeed. Prior to Summer 2020, to remove the hold, students were required to take the SCORE (Student Core Outcomes and Readiness Evaluation) at Testing Services, located on the second floor of Biggin Hall. If students were off-campus during their last semester they were administered the SCORE in a remote proctored environment (Proctor U; 228 students). However, in March of 2020, Auburn University moved all instruction to the remote environment due to the COVID-19 Pandemic. A small group of students (124) were administered the SCORE remotely (no proctor). Then in Summer 2020, all assessments were administered remotely (767 students). In Fall of 2020, students were permitted to take the assessment on-campus or remotely. A handful of students selected to take the test at Test Services (143 students) while most completed the assessment remotely (1072 students). Of note, 67 students completed the

assessment during the Spring but then graduated during a later semester (59 on-campus, 2 remote-proctored, 6 remote). In addition, a small portion of Waivers were permitted due to technical difficulties. Overall, 95.4% of Spring graduates completed the SCORE; 94.7% of Summer graduates completed the SCORE; 89.0% of Fall graduates completed the SCORE. Failure to remove the AT Hold through testing or a waiver results in the student not receiving their diploma (though the degree will still be conferred, pending successful fulfillment of other degree requirements).

PARTICIPATION

A total of 5151 students took the ETS HEIghten Critical Thinking Assessment. Testing occurred over the Spring, Summer, and Fall of 2020; 3102 students took the test in the Spring, 834 students took it in the Summer, and the remaining test takers were graduating seniors in the Fall. Of note, there were no meaningful differences in critical thinking scores between students taking the test in the Spring, Summer, or Fall (see Table 6) although performance was, on average, slightly higher for Spring graduates. This allows for a collapse of the data across the three cohorts for further analysis.

In total, 3252 students graduated in the Spring, 881 students graduated in the Summer, and 1365 students graduated in the Fall semester for a total of 5498 students. Table 1 compares the present sample to the graduating population. Of note, the percentage of "Legacy" students includes students who had a parent, sibling, or "Other" individual attend Auburn University.

Figure 2 displays the proportion of test-takers by College relative to the overall population of Spring, Summer, and Fall 2020 graduates. The present sample is very similar to the population.

Table 2 shows differences in student demographics across 2020 cohorts. Cohorts are generally similar; however, the Fall cohort consisted of more male, less white, more first generation, and less legacy students. The Summer cohort along with Fall cohort had more transfer students and, on average, lower ACT scores and GPA. Closely related to this are the demographic differences between the on-campus and remote test takers displayed in Table 3 (given that 93% of on-campus test takers were Spring 2020 graduates).

Table 1 – Demographic Data on All Seniors and the SCORE Sample

	N	% Male	% White	%Transfer	% First Gen	% Legacy	ACT Avg (SD)	GPA Avg (SD)
Spring, Summer and Fall 2020 Graduates	5498	50.8%	79.4%	21.5%	15.2%	42.8%	26.5 (4.2)	3.23 (0.5)
Critical Thinking Test-Takers	5151	50.5%	79.6%	21.4%	15%	42.7%	26.6 (4.2)	3.25 (0.5)

Figure 2 – Senior Population and Testing Sample by College

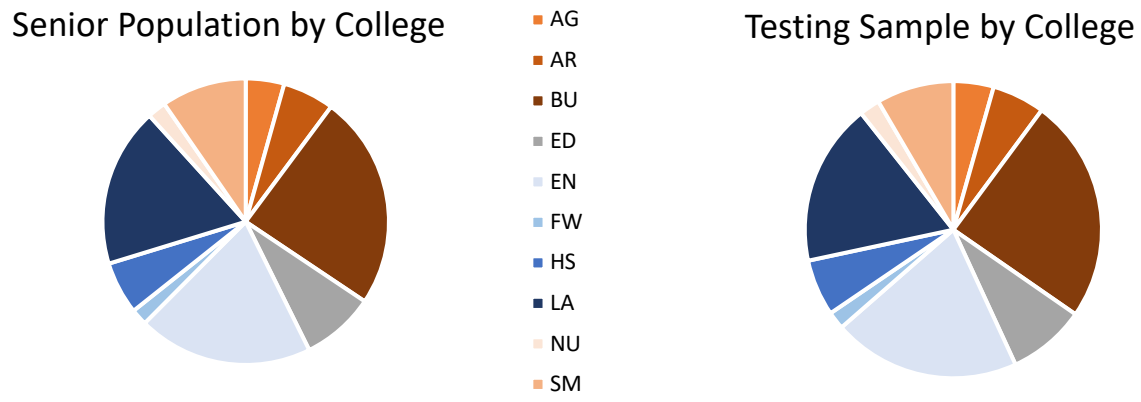


Table 2 – Demographic Data by Cohort

	N Graduates	N Test Takers	% Male	% White	% Transfer	% First Gen	% Legacy	ACT Avg (SD)	GPA Avg (SD)	% Remote Test-Takers
Spring 2020	3252	3102	49.4%	81.7%	18.1%	15.0%	43.0%	27.1 (4.2)	3.33 (0.5)	4.0%* 7.4%**
Summer 2020	881	834	48.3%	80.6%	24.9%	13.2%	44.2%	25.5 (3.9)	3.11 (0.5)	92.9%***
Fall 2020	1365	1215	55.0%	73.7%	27.2%	16.4%	40.9%	25.8 (4.1)	3.15 (0.5)	88.2%

**Not proctored; **Proctored;*

****67 students took the critical thinking test in Spring 2020 but graduated in Summer 2020 (59 out of 67 took it on campus)*

Table 3 – Demographic Data by Test-Taking Setting

	N	% Male	% White	%Transfer	% First Gen	% Legacy	ACT Avg (SD)	GPA Avg (SD)
On Campus	2952	51.8%	80.5%	18.6%	15.2%	41.8%	27.1 (4.2)	3.3 (0.5)
Remote*	2199	48.8%	78.4%	25.1%	14.7%	44%	25.8 (4.1)	3.2 (0.5)

**230 students (10.5%) – Remote proctored*

Of Note: Almost all on-campus test-takers were from Spring 2020 Cohort (93%; 2750 out of 2952 on-campus test takers)

STUDENT MOTIVATION

When presented with a nonconsequential assessment, students may approach the task differently. Some students may put forth their best effort. Others may try, but not as hard as they would if the test impacted them personally. Others still may not try at all and rapidly respond to test items without giving them consideration (Wise & DeMars, 2005). This variability is problematic because if students do not put forth their best effort on an assessment, their score will underestimate their ability (limiting the validity of score interpretations). Thus, students' test-taking motivation was assessed to determine which responses may lack validity. Data that clearly do not reflect students' best effort on the assessment were deleted. The process used to determine the cases to be deleted is outlined below.

IDENTIFYING AND DELETING 'UNMOTIVATED' CASES

The sample data were carefully evaluated to flag cases where students clearly did not put forth their best effort. This section of the report outlines characteristics of the full dataset and decision points to delete students classified as 'unmotivated.'

SELF-REPORTED EFFORT

The ETS HElghten Critical Thinking Assessment includes a question on student effort (as it related to effort put forth during the test). Students responded to the effort question using a 3-point scale with options being: "No effort", "I gave some effort", and "I tried my best."

As expected, the "No effort" group, on average, had lower critical thinking scores and spent less time on the test (i.e., the median time spent on the test was 13% of the allowed time). 560 students reported putting forth "no effort" on the test and were excluded from further analyses. Given the possibility of unmotivated students responding carelessly to the ETS effort question, test time

estimates were further evaluated for the remaining students. Out of the remaining 4591 students, 230 spent 13% of the allowed time or less on the test indicating careless responding and were thus excluded (Of note: 13% was the median percentage of time spent on the test for the group that put forth no effort). Table 4 displays critical thinking test scores, average test time used, and additional demographics of students identified as unmotivated.

Table 4 – Unmotivated Students

		Critical Thinking Test Score			Time Used				
	N	Average	Min	Max	Average	SD	% On-campus test-takers	% ETS 'Advanced' Classification	% Male
Indicated 'No Effort'	560	155.1	150	177	21.7%	22.8	54.1%	1.4%	62%
Remaining Low Time Estimates	230	152.8	150	167	7.4%	3.9	43.9%	0%	58.3%

CRITICAL THINKING SCORES

Students receiving the lowest possible value on all three scores (i.e., overall, analytic, and synthetic) were removed. This resulted in the removal of 33 cases and a final sample size of N = 4,328.

THE TEST

The ETS HEIghten Critical Thinking assessment provides an Overall Critical Thinking score which ranges from 150-180. ETS has also created three Proficiency scores: "Developing" (for those students scoring between 150-161), "Proficient" (for those students scoring between 162-172), and "Advanced" (for those students scoring between 173-180). See Appendix A for detailed information about the characteristics of students in each classification area.

Additionally, the ETS produces sub-scores for two areas: Analytical and Synthetic dimensions. These scores are reported on a 1-10 scale, with 10 being the highest score.

- **The Analytic Dimension** assesses a student's ability to 1) evaluate evidence and its use, and 2) analyze and evaluate arguments.
- **The Synthetic dimension** measures a student's ability to 1) understand implications and consequences and 2) develop sound and valid arguments. As a reminder, our Sub-outcome 1 is aligned with the HEIgthen tests' "Analytic" score.

TEST PERFORMANCE

Table 5 provides a summary of student performance on the Critical Thinking assessment. Of note, these data do not include the 790 'unmotivated' or 33 'lowest score possible' cases.

Table 5 – Critical Thinking Scores

Sub-Outcome	HEIghten Area	Average	SD	Min	Max
Identify the genre of the text, make reasonable inferences about its central purpose or argument, define its key components, and show how the writer uses these to reach a conclusion or create meaning or impact.	Sub-score- Analytic Dimension	5.0	2.4	1.0	10
Construct a strong, well-reasoned argument by determining which conclusion is supported by the strongest evidence.	Sub-score- Synthetic Dimension	5.1	2.2	1.0	10
Engage the text dialogically, questioning its premises, identifying its limitations, or advancing alternative perspective.	Overall Critical Thinking	165.3	6.9	150	180

N = 4,328

Figure 3 is a visual representation of the number of students in each HEIghten Proficiency category. For more information on each of the Proficiency categories, see Appendix A.

Table 6 compares student performance on the Critical Thinking assessment across three 2020 cohorts. There were no meaningful differences in scores across 2020 cohorts, which will allow us to collapse the data in further analyses.

Figure 3 – Critical Thinking 2020 Proficiency Scores

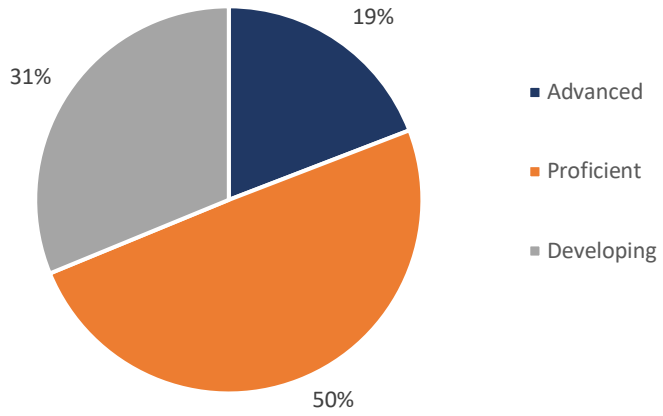


Table 6 – Critical Thinking Scores by Cohort

	N	HEIghten Area	Average	SD	Min	Max
Spring 2020	2643	Sub-score- Analytic Dimension	5.1	2.4	1.0	10
		Sub-score- Synthetic Dimension	5.2	2.2	1.0	10
		Overall Critical Thinking	165.8	7.0	150	180
Summer 2020	667	Sub-score- Analytic Dimension	4.7	2.2	1.0	9.8
		Sub-score- Synthetic Dimension	4.8	2.1	1.0	9.8
		Overall Critical Thinking	164.3	6.6	150	180
Fall 2020	1018	Sub-score- Analytic Dimension	4.7	2.3	1.0	10
		Sub-score- Synthetic Dimension	4.9	2.3	1.0	10
		Overall Critical Thinking	164.6	6.9	150	180

Table 7 compares student performance on the Critical Thinking assessment between 2018 and 2020 Summer/Fall graduates. There was no meaningful change in critical thinking scores between Summer/Fall 2018 and Summer/Fall 2020 cohorts. Given that no critical thinking intervention was implemented at Auburn University in the meantime, these results confirm consistency of the ETS Critical Thinking assessment. Additionally, performance scores are almost the same despite the 2018 sample being much smaller than the 2020 sample, which confirms the quality of random sampling of students in 2018.

Table 7 – Critical Thinking Scores in Summer and Fall of 2018 vs. Summer and Fall of 2020

	N	HEIghten Area	Average	SD	Min	Max
2018	501	Sub-score- Analytic Dimension	4.7	2.3	1.0	10
		Sub-score- Synthetic Dimension	4.7	2.4	1.0	10
		Overall Critical Thinking	164.1	7.3	150	180
2020	1685	Sub-score- Analytic Dimension	4.7	2.3	1.0	9.8
		Sub-score- Synthetic Dimension	4.9	2.2	1.0	9.8
		Overall Critical Thinking	164.5	6.8	150	180

AUBURN SAMPLE COMPARED TO ETS SAMPLE

A total of 25,936 students took the Critical Thinking assessment in the ETS sample. The average total score for this group of students was **162.8**. Likewise, the average analytic score for these students was 4.0 while the average synthetic score was 4.0. Thus, it appears that the Auburn student sample performed better than the normed group, (overall - **164.1**; analytic - 4.6; synthetic - 4.8). In addition, based on the ETS proficiency classification, 66% of Auburn students performed in Proficient and Advanced categories compared to 55% of students in the comparison group. Of note, the ETS sample likely included students at all academic levels (i.e., freshmen, sophomores, juniors, and seniors), and the Auburn sample was composed only of seniors.

NOTEWORTHY TEST RELATIONSHIPS

The Critical Thinking test can be associated with internal data from Auburn to explore critical thinking skills across certain demographics and situations. First, test scores were correlated with ACT scores as one may hypothesize that some students are better test-takers than others. Second, test scores were disaggregated by grades earned in the following courses: ENGL 2200/2210 (World Literature before 1600 & after 1600); ENGL 2230/2240 (British Literature before 1789 & after 1789; ENGL 2250/2260 (American Literature before 1865 & after 1865); PHIL 1020 – Introduction to Ethics; PHIL 1040 – Business Ethics. Additional philosophy courses are options in the core curriculum; however, these courses had low enrollments and thus these data were not analyzed. It is anticipated that higher grades earned in these courses would result in higher test scores, providing some validity evidence to test score interpretation.

ACT CONVERTED SCORES

Pearson correlations were conducted to determine the strength of the relationship between students' ACT scores and their critical thinking scores; results are displayed in Table 8. Of note, ACT scores are statistically significantly related to all critical thinking scores, with the strongest relationship between ACT scores and the Overall score (r = 0.59).

Table 8 – Critical Thinking and ACT Scores

	N	ACT Scores
Overall	3,780	0.59***
Analytic	3,780	0.51***
Synthetic	3,780	0.54***

****p* < .001
Of Note: Only 3,780 converted ACT scores were available in the Banner dataset.

COURSE GRADES

ENGL 2200 (WORLD LITERATURE BEFORE 1600) GRADES

****Note: Students receiving grades of D, F, W were excluded in the following analyses due to small sample sizes.*

Table 9 displays descriptive statistics for students earning an A, B, or C in ENGL 2220. Students' average Critical Thinking scores were consistently higher for higher grades earned in this course.

Table 9 – ENGL 2220 Grades

	N	Overall	Analytic	Synthetic
ENGL 2220 - A	194	168.4 (6.9)	6.0 (2.4)	5.9 (2.2)
ENGL 2220 - B	173	165.2 (6.9)	5.0 (2.3)	5.0 (2.2)
ENGL 2220 - C	89	163.0 (6.4)	4.3 (2.3)	4.4 (2.0)

ENGL 2210 (WORLD LITERATURE AFTER 1600) GRADES

Table 10 displays descriptive statistics for students earning an A, B or C in ENGL 2210. Students' average Critical Thinking scores were consistently higher for higher grades earned in this course.

Table 10 – ENGL 2210 Grades

	N	Overall	Analytic	Synthetic
ENGL 2210 - A	224	167.1 (7.1)	5.5 (2.4)	5.7 (2.3)
ENGL 2210 - B	160	164.7 (6.4)	4.7 (2.3)	5.0 (2.0)
ENGL 2210 - C	50	162.8 (5.4)	4.4 (2.2)	4.2 (1.8)

ENGL 2230 (BRITISH LITERATURE BEFORE 1789) GRADES

Table 11 displays descriptive statistics for students earning an A or B in ENGL 2230. Students' average Critical Thinking scores were consistently higher for higher grades earned in this course.

Table 11 – ENGL 2230 Grades

	N	Overall	Analytic	Synthetic
ENGL 2230 - A	231	167.3 (6.3)	5.6 (2.3)	5.7 (2.1)
ENGL 2230 - B	153	164.4 (6.1)	4.7 (2.2)	4.9 (2.1)
ENGL 2230 - C	44	161.3 (6.7)	3.9 (2.2)	3.9 (1.97)

ENGL 2240 (BRITISH LITERATURE AFTER 1789) GRADES

Table 12 displays descriptive statistics for students earning an A or B in ENGL 2240. Students' average Critical Thinking scores were consistently higher for higher grades earned in this course.

Table 12 – ENGL 2240 Grades

	N	Overall	Analytic	Synthetic
ENGL 2240 - A	270	168.5 (6.5)	5.9 (2.3)	6.1 (2.2)
ENGL 2240 - B	187	165.1 (7.1)	4.9 (2.3)	5.0 (2.2)
ENGL 2240 - C	50	162.6 (7.5)	4.3 (2.6)	4.2 (2.2)

ENGL 2250 (AMERICAN LITERATURE BEFORE 1865) GRADES

Table 13 displays descriptive statistics for students earning an A, B, or C in ENGL 2250. Students' average Critical Thinking scores were consistently higher for higher grades earned in this course.

Table 13 – ENGL 2250 Grades

	N	Overall	Analytic	Synthetic
ENGL 2250 - A	256	166.97 (6.6)	5.6 (2.3)	5.5 (2.2)
ENGL 2250 – B	198	163.3 (6.5)	4.4 (2.2)	4.5 (2.2)
ENGL 2250 – C	53	160.6 (6.6)	3.4 (1.96)	3.97 (2.3)

ENGL 2260 (AMERICAN LITERATURE AFTER 1865) GRADES

Table 14 displays descriptive statistics for students earning an A, B, or C in ENGL 2260. Students' average Critical Thinking scores were consistently higher for higher grades earned in this course.

Table 14 – ENGL 2260 Grades

	N	Overall	Analytic	Synthetic
ENGL 2260 - A	217	168.6 (6.5)	5.9 (2.2)	6.1 (2.2)
ENGL 2260 – B	205	164.7 (6.4)	4.8 (2.2)	4.9 (2.1)
ENGL 2260 – C	66	162.9 (6.1)	4.3 (2.1)	4.4 (2.1)

PHIL 1020 (INTRODUCTION TO ETHICS) GRADES

Table 15 displays descriptive statistics for students earning an A, B, or C in PHIL 1020. Students' average Critical Thinking scores were consistently higher for higher grades earned in this course.

Table 15 – PHIL 1020 Grades

	N	Overall	Analytic	Synthetic
PHIL 1020 – A	524	167.9 (6.6)	5.7 (2.3)	5.9 (2.2)
PHIL 1020 – B	470	164.3 (6.8)	4.7 (2.3)	4.8 (2.2)
PHIL 1020 – C	188	162.4 (6.5)	4.0 (2.1)	4.3 (2.1)

PHIL 1040 (BUSINESS ETHICS) GRADES

Table 16 displays descriptive statistics for students earning an A, B, or C in PHIL 1040. Students' average Critical Thinking scores were consistently higher for higher grades earned in this course.

Table 16 – PHIL 1040 Grades

	N	Overall	Analytic	Synthetic
PHIL 1040 – A	446	166.95 (7.0)	5.4 (2.4)	5.6 (2.3)
PHIL 1040 – B	268	162.8 (6.3)	4.3 (2.2)	4.4 (2.0)
PHIL 1040 – C	63	161.5 (6.7)	3.8 (2.3)	4.01 (2.1)

Generally, students who got higher grades in core courses (aligned with the critical thinking outcome) scored higher on the ETS Critical Thinking assessment. Further investigation indicated that ACT scores explain 35% (R^2=.3507) of unique variance in Critical Thinking scores above and beyond grades in core courses. When added to the model, grades for most core courses added less than 1% to the variance explained ($\Delta R^2 < .01$). However, ENGL 2200 explained additional 4.13% of variance ($\Delta R^2 = .0413$), ENGL 2260 added 2.87% ($\Delta R^2 = .0287$) while ENGL 2250 contributed with 1.29% of unique variance explained in Overall Critical Thinking scores ($\Delta R^2 = .0129$). Thus, although ACT scores are a much better predictor of performance on Critical Thinking assessment than any core course grades, some core courses make a small but unique contribution to the prediction of performance.

EXPLORATORY ANALYSES

In the Summer of 2018, Dr. Megan Good (Director, Academic Quality and Improvement) invited internal campus stakeholders to create a list of research questions to apply to the SCORE data. Representatives from English, Philosophy, Institutional Research, Accessibility, Student Affairs, the Library, and Academic Support were in attendance and generated the research questions below based on Banner data, which fall into two categories: Demographics and Curricular Experiences. The research questions were generated to explore areas where learning gaps may exist. Appendix B includes additional learning research questions that could be explored if analytic capacity were available.

DEMOGRAPHICS

Campus stakeholders were interested in whether or not sub-populations of students scored differently based on demographic characteristics. Below are the results related to demographic differences by gender, race, first-generation, legacy, and transfer status.

GENDER

Table 17 displays the average score for males and females across critical thinking scores. Independent sample t-tests showed statistically significant differences in overall, analytic, and synthetic scores between these groups. However, given the negligible effect sizes, the statistical significance was likely the result of the test being overpowered with the large sample sizes:

- Males had greater Overall Critical Thinking scores (M=165.6, SD = 7.2) than females (M=164.9, SD = 6.6). This difference was statistically significant, t (4326) = 3.55, p = .0004, but with a negligible effect size (Cohen's d = .11).
- Males had greater Analytic scores (M=5.1, SD = 2.4) than females (M=4.8, SD = 2.3). This difference was statistically significant, t (4326) = 3.42, p = .0006, but with a negligible effect size (Cohen's d = .10).
- Males had greater Synthetic scores (M=5.2, SD = 2.3) females (M=5.0, SD = 2.2). This difference was statistically significant, t (4326) = 3.26, p = .001, but with a negligible effect size (Cohen's d = .099).

Thus, there were no meaningful differences in Critical Thinking scores between males and females.

Table 17 – Gender Differences

	N	Overall	Analytic	Synthetic
Male	2105	165.6 (7.2)	5.1 (2.4)	5.2 (2.3)
Female	2223	164.9 (6.6)	4.8 (2.3)	5.0 (2.2)

RACE

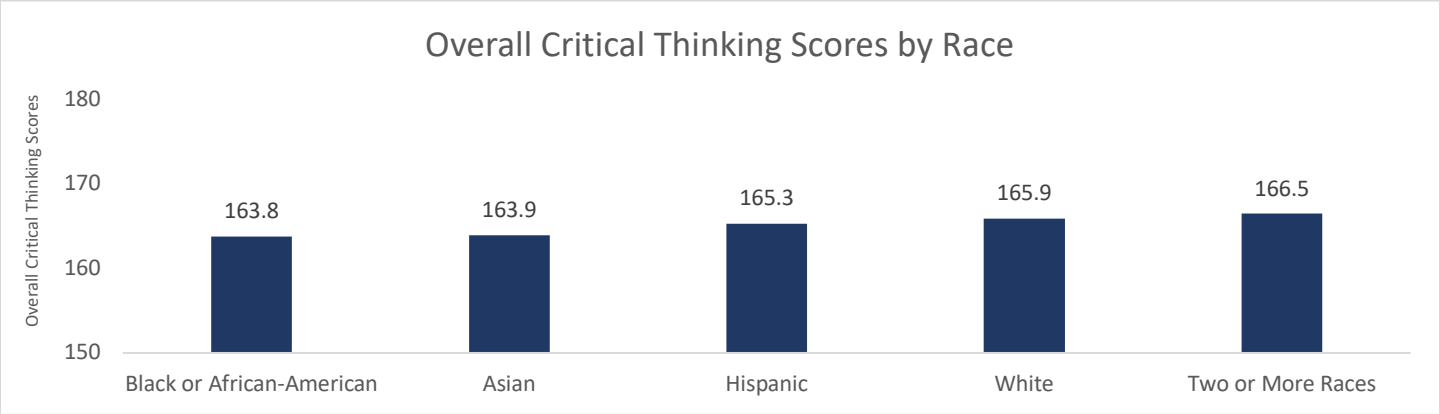
Table 18 and Figure 4 display the average scores for different races across the critical thinking scores. Of note, statistical significance was not explored given the unbalanced sample sizes.

Table 18 – Racial Differences

	N	Overall	Analytic	Synthetic
White	3426	165.9 (6.8)	5.2 (2.3)	5.3 (2.2)
Non-White	599	164.6 (6.7)	4.8 (2.2)	4.9 (2.2)

Of Note: Non-White group consists of American Indian or Alaska Native, Native Hawaiian/Pacific Isl., Hispanic, Asian, Black or African-American, Two or More Races, and Unknown; non-resident aliens excluded.

Figure 4 – Overall Average Critical Thinking Scores by Race (subgroups with sample sizes smaller than 15 were excluded)



FIRST GENERATION

Table 19 displays the average score for students tagged as "First Generation College Students" and those who are not. Of note, statistical significance was not explored given the unbalanced sample sizes.

Table 19 – First Generation

	N	Overall	Analytic	Synthetic
NOT First Generation	3665	165.3 (7.0)	5.0 (2.4)	5.1 (2.2)
First Generation	663	164.8 (6.7)	4.9 (2.3)	4.9 (2.2)

LEGACY

****Legacy student is defined as: student having a "Parent", "Sibling," or "Other" attend Auburn.*

Table 20 displays the average scores for students who identify as Legacy (i.e., a parent, sibling, or "Other" person attended Auburn) and those who did not. Table 21 groups together the three Legacy categories. Independent sample t-tests showed statistically significant differences in overall, analytic, and synthetic scores between these groups. However, given the negligible effect sizes, the statistical significance was likely the result of the test being overpowered with the large sample sizes:

- Students who identify as legacy had greater Overall Critical Thinking scores (M=165.6, SD = 6.8) than students who do not identify as legacy (M=165.0, SD = 7.0). This difference was statistically significant, t (4326) = 3.05, p = .002, but with a negligible effect size (Cohen's d = .09).
- Students who identify as legacy had greater Analytic scores (M=5.1, SD = 2.3) than students who do not identify as legacy (M=4.9, SD = 2.4). This difference was statistically significant, t (4326) = 2.1, p = .04, but with a negligible effect size (Cohen's d = .10).
- Students who identify as legacy had greater Synthetic scores (M=5.2, SD = 2.2) than students who do not identify as legacy (M=5.0, SD = 2.2). This difference was statistically significant, t (4326) = 3.34, p = .0008, but with a negligible effect size (Cohen's d = .06).

Thus, there were no meaningful differences in Critical Thinking scores between students who identify as legacy and students who do not identify as legacy at Auburn University.

Table 20 – Legacy

	N	Overall	Analytic	Synthetic
Student reported "Other"	599	165.6 (6.5)	5.1 (2.2)	5.1 (2.1)
Student reported "Parent"	1047	165.8 (7.0)	5.1 (2.4)	5.3 (2.3)
Student reported "Sibling"	163	165.1 (7.3)	4.9 (2.4)	5.0 (2.4)
No Affiliation	2519	165.0 (7.0)	4.9 (2.4)	5.0 (2.2)

Table 21 – Legacy: Grouped

	N	Overall	Analytic	Synthetic
Student reported "Other", "Parent" or "Sibling"	1809	165.6 (6.8)	5.1 (2.3)	5.2 (2.2)
No Affiliation	2519	165.0 (7.0)	4.9 (2.4)	5.0 (2.2)

TRANSFER

Table 22 displays the average score for transfer students and those who started their Auburn career as freshmen. Of note, students who began their Auburn career as freshmen performed statistically significantly higher on all critical thinking domains than students who transferred in, specifically:

- Students who started at Auburn as freshmen had greater Overall Critical Thinking scores (M=166.5, SD = 6.8) than students who transferred in (M=162.9, SD = 6.1). This difference was statistically significantly different, t (4077) = 14.59, $p < 0.001$, with a moderate effect size (Cohen's d = .55).
- Students who started at Auburn as freshmen had greater Analytic scores (M=5.3, SD = 2.3) than students who transferred in (M=4.3, SD = 2.1). This difference was statistically significantly different, t (4077) = 12.24, $p < 0.001$, with a moderate effect size (Cohen's d = 0.46).
- Students who started at Auburn as freshmen had greater Synthetic scores (M=5.5, SD = 2.2) than students who transferred in (M=4.3, SD = 2.1). This difference was statistically significantly different, t (4077) = 13.61, $p < 0.001$, with a moderate effect size (Cohen's d = 0.51).

Table 22 – Transfer

	N	Overall	Analytic	Synthetic
Transfer	926	162.9 (6.1)	4.3 (2.1)	4.3 (2.1)
Non-Transfer (Started as Freshmen)	3153	166.5 (6.8)	5.3 (2.3)	5.5 (2.2)

We suspected that the differences in Critical Thinking scores between transfer and non-transfer students might be attributed to differences in their ACT scores. A sequential regression demonstrated that ACT scores explain 35% (R^2=.3507) of unique variance in Critical Thinking scores above and beyond transfer status. When added to the model, transfer status (ΔR^2=.0078) only added .8% to the variance explained. Thus, ACT scores are a much better predictor of performance on Critical Thinking assessment than transfer status.

CURRICULAR EXPERIENCES

Campus stakeholders were interested in how various aspects of the Auburn experience related to Critical Thinking scores. Below are the results related to student engagement with internships (as measured in CEES), College, and the Core Curriculum Sequence requirement.

INTERNSHIPS

Table 23 displays descriptive statistics for these two groups. Independent sample t-tests showed statistically significant differences in overall and synthetic scores between these groups. However, given the negligible effect sizes, the statistical significance was likely the result of the test being overpowered with the large sample sizes:

- Students who did not participate in internship had greater Overall Critical Thinking scores (*M*=165.5, *SD* = 6.9) than students who participated in internship (*M*=165.0, *SD* = 6.9). This difference was statistically significant, *t* (4307) = 2.36, *p* = .02, but with a negligible effect size (Cohen's *d* = .07).
- Students who did not participate in internship had greater Synthetic scores (*M*=5.2, *SD* = 2.2) than students who participated in internship (*M*=5.0, *SD* = 2.2). This difference was statistically significant, *t* (4307) = 2.55, *p* = .01, but with a negligible effect size (Cohen's *d* = .08).

Thus, there were no meaningful differences in Critical Thinking scores between students who participated in internship and students who do not did not participate in internship at Auburn.

Table 23 – Internships

	N	Overall	Analytic	Synthetic
Internship	2081	165.0 (6.9)	4.9 (2.3)	5.0 (2.2)
No Internship	2247	165.5 (6.9)	5.0 (2.4)	5.2 (2.2)

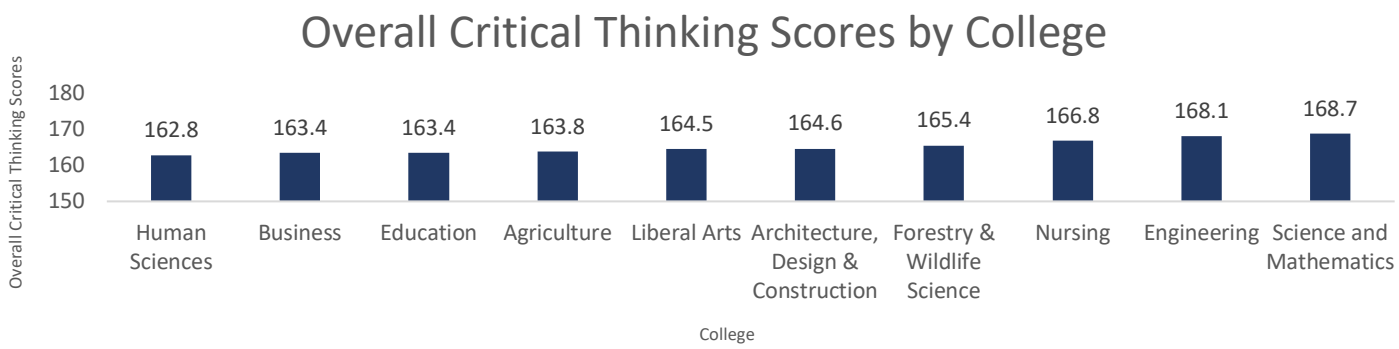
COLLEGE

Table 24 and Figure 5 display descriptive statistics for each College. Of note, students from the College of Science and Mathematics had the highest, average Overall Critical Thinking scores (*M* = 168.7, *SD* = 5.6) followed by the College of Engineering (*M* = 168.1, *SD* = 6.9) whereas students from the College of Human Sciences had the lowest, average Overall Critical Thinking scores (*M* = 162.8, *SD* = 6.3).

Table 24 – College Differences

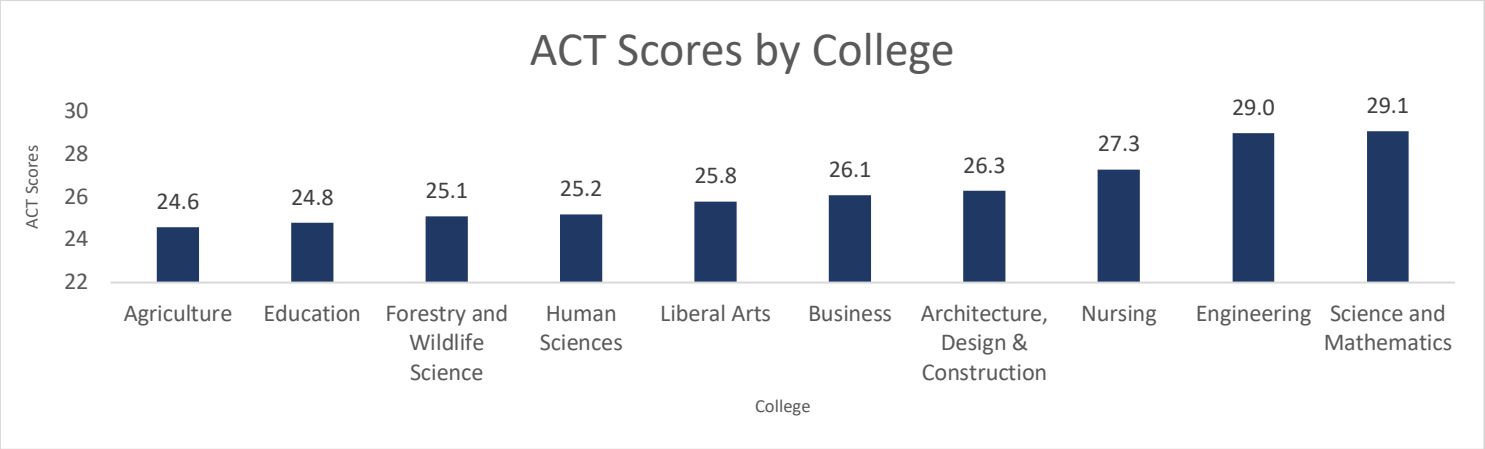
	N	Overall	Analytic	Synthetic
Agriculture	183	163.8 (6.5)	4.4 (2.3)	4.7 (2.1)
Architecture, Design & Construction	235	164.6 (6.7)	4.8 (2.4)	4.9 (2.1)
Business	993	163.4 (7.0)	4.4 (2.4)	4.5 (2.2)
Education	367	163.4 (5.9)	4.5 (2.1)	4.5 (2.0)
Engineering	932	168.1 (6.9)	5.7 (2.4)	6.0 (2.2)
Forestry & Wildlife Science	82	165.4 (7.0)	5.2 (2.3)	5.0 (2.1)
Human Sciences	273	162.8 (6.3)	4.2 (2.1)	4.4 (2.1)
Liberal Arts	764	164.5 (6.7)	4.7 (2.3)	4.9 (2.2)
Nursing	104	166.8 (6.2)	5.4 (2.2)	5.6 (2.0)
Science and Mathematics	395	168.7 (5.6)	6.0 (2.1)	6.1 (1.9)

Figure 5 – Critical Thinking Scores by College



Given the high correlation between Overall Critical Thinking scores and ACT scores, we further explored ACT scores by college. Figure 6 displays average ACT scores for each College. The scores showed the similar pattern with the College of Science and Mathematics having the highest, average ACT scores (*M* = 29.1, *SD* = 3.7) followed by the College of Engineering (*M* = 29.0, *SD* = 3.9). On the other hand, the College of Human Sciences had one of the lowest, average ACT scores (*M* = 25.2, *SD* = 3.6). Thus, ACT scores are once again a better predictor of performance on Critical Thinking assessment than College.

Figure 6 – ACT Scores by College



CORE CURRICULUM SEQUENCE – AUBURN ONLY OR NOT

Table 25 displays descriptive statistics for students earning credits for **both** parts of a sequence in either literature or history from Auburn University and for students who took part of the sequence or whole sequence elsewhere (e.g., AP credit, transfer credit). Independent sample t-tests showed statistically significant differences in overall, analytic, and synthetic scores between these groups. However, given the negligible effect sizes, the statistical significance was likely the result of the test being overpowered with the large sample sizes:

- Students who took both Core Curriculum Sequence classes at Auburn had greater Overall Critical Thinking scores (*M*=165.6, *SD* = 6.9) than students who transferred in some credit for the Core Curriculum Sequence (*M*=164.8, *SD* = 6.9). This difference was statistically significant, t (4326) = 3.9, $p < .001$, but with a negligible effect size (Cohen's d = .12).
- Students who took both Core Curriculum Sequence classes at Auburn had greater Analytic scores (*M*=5.0, *SD* = 2.4) than students who transferred in some credit for the Core Curriculum Sequence (*M*=4.8, *SD* = 2.3). This difference was statistically significant, t (4326) = 2.9, p = .004, but with a negligible effect size (Cohen's d = .09).

- Students who took both Core Curriculum Sequence classes at Auburn had greater Synthetic scores (M=5.2, SD = 2.2) than students who transferred in some credit for the Core Curriculum Sequence (M=4.9, SD = 2.2). This difference was statistically significant, t (4326) = 4.1, p < .001, but with a negligible effect size (Cohen's d = .12).

Thus, there were no meaningful differences in Critical Thinking scores between students who took both Core Curriculum Sequence classes at Auburn and students who transferred in some credit for the Core Curriculum Sequence.

Table 25 – Core Sequence Credit Earned at Auburn or Not

	N	Overall	Analytic	Synthetic
Both Sequence Courses Taken at Auburn	2390	165.6 (6.9)	5.0 (2.4)	5.2 (2.2)
Both Sequence Courses **NOT** Taken at AU	1938	164.8 (6.9)	4.8 (2.3)	4.9 (2.2)

CORE CURRICULUM SEQUENCE – CHOICE

Table 26 and Figure 7 display descriptive statistics for students' choice of sequence of the 2,390 students who took both sequence courses at Auburn. Of note, 128 students took two sequences (e.g., American Literature and World History). Statistical significance testing was not conducted given the unbalanced sample sizes with most students taking the World History sequence.

Table 26 – Core Sequence Choice

	N	Overall	Analytic	Synthetic
Two or More Sequences Taken	128	166.7 (7.1)	5.4 (2.5)	5.5 (2.3)
British Literature Sequence	92	166.2 (7.1)	5.3 (2.6)	5.3 (2.1)
World Literature Sequence	90	167.4 (6.5)	5.6 (2.3)	5.7 (2.1)
American Literature Sequence	91	164.6 (6.5)	4.7 (2.2)	5.0 (2.2)
World History Sequence	1815	165.3 (6.9)	4.9 (2.4)	5.1 (2.2)
Tech and Civ Sequence	174	167.7 (6.9)	5.7 (2.4)	5.8 (2.2)

Figure 7 – Critical Thinking Scores by Auburn Sequence Chosen

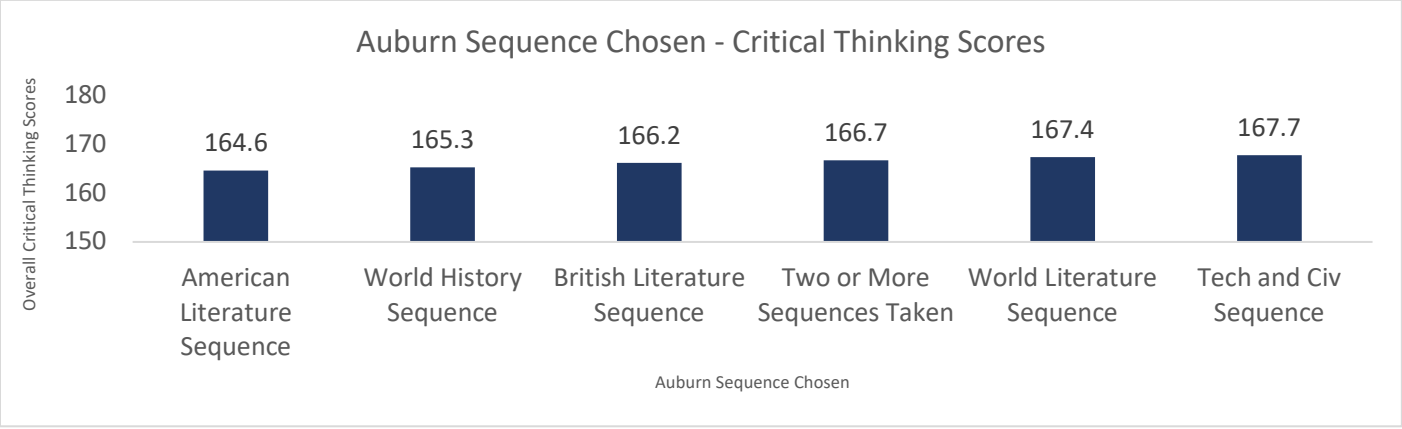
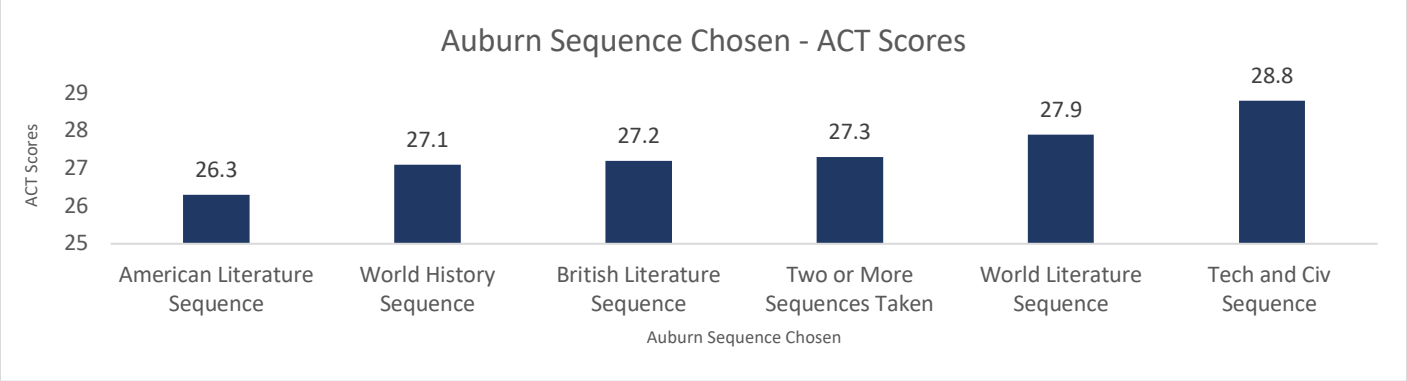


Figure 8 shows average ACT scores by students' choice of core curriculum sequence. Given the same pattern of average ACT and Critical Thinking scores across core curriculum sequences, we can conclude that the differences in ACT scores are once again responsible for the observed differences in Critical Thinking scores.

Figure 8 – ACT Scores by Auburn Sequence Chosen



MAIN TAKEAWAYS

- Auburn students outperformed the ETS peer group on the Critical Thinking assessment in Overall, Analytic, and Synthetic scores.
- ACT scores explain 35% of unique variance in performance on Critical Thinking assessment and are thus the best predictor of Critical Thinking scores. This is likely because both ACT and Critical Thinking test are standardized tests that are greatly affected by general cognitive ability.
- When it comes to core courses aligned with critical thinking, generally, students who got higher grades in core courses scored higher on the ETS Critical Thinking assessment. Although this is again related to ACT scores, performance in one of the following courses: ENGL 2200, ENGL 2250, and ENGL 2260 also predicts performance on the Critical Thinking test.
- Gaps in knowledge were only identified among transfer students. In other words, students who began their higher education career at Auburn as freshmen outperformed transfer students on all aspects of the Critical Thinking test. However, further analysis showed that these differences were due to differences in ACT scores. Thus, potential critical thinking interventions should target all students with low ACT scores and not just transfer students.

REFERENCES

Educational Testing Service (ETS; 2017). HEIghten Outcomes Assessment Suite: Guide to Score Interpretation.

Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational assessment*, *10*(1), 1-17.

Table 6. Critical Thinking

Advanced	Proficient	Developing
A typical student at the advanced level has demonstrated the ability to:	*A typical student at the proficient level has demonstrated the ability to:*	*A typical student at the developing level may sometimes:*
extrapolate implications from multiple pieces of information and argumentation	make inferential connections between points whose relationship is not explicitly given	make inferential connections between two explicitly related points
accurately recognize descriptions of the logic of complexly structured arguments	follow the logic of an argument whose structure is not fully explicit	follow the logic of an explicitly structured argument
employ multi-step reasoning to identify hidden assumptions	identify implicit assumptions	identify explicit assumptions
employ multi-step reasoning to identify evidence that directly or indirectly supports or undermines a claim, or specify additional information needed in order to resolve a point	identify evidence that directly or indirectly supports or undermines a claim or specify additional information needed in order to resolve a point	identify evidence that directly supports or undermines a claim
identify subtle appeals to emotion and revisions to an argument that would reduce such appeals	identify appeals to emotion and revisions to an argument that would reduce such appeals	identify clear appeals to emotion
distinguish information that may be peripherally or generally relevant to assertions/arguments from information that is directly on-point	distinguish information that is relevant to assertions or arguments from irrelevant information	mistake evidence that is broadly related to a topic for evidence that is relevant to a specific assertion about the topic
employ multi-step reasoning to distinguish causation from correlation, and identify possible alternative causes or explanations	distinguish causation from correlation, and identify possible alternative causes or explanations	have difficulty distinguishing causation from correlation or identifying alternative explanations
engage in reasoning that involves complex interactions among multiple claims, arguments or pieces of information	engage in reasoning that involves interactions among multiple claims, arguments or pieces of information	have difficulty understanding or evaluating interactions among multiple claims,

APPENDIX B – RESEARCH QUESTIONS FOR FUTURE CONSIDERATION

Below are learning research questions that expand beyond Banner data that could be answered if additional analytic capacity were available. Of note, this list is not exhaustive; SCORE data can be connected to any dataset that includes student IDs or GIDs.

- Co-Curricular
    - Is there a relationship between critical thinking skills and the number of Service Hours students expend (Data Source: AU Involve)?
    - Is there a relationship between critical thinking skills and the number of active memberships students have with student organizations (Data Source: AU Involve)?
    - Is there a relationship between critical thinking skills and student involvement in Greek life (Data Source: AU Involve)?
        - Does the above relationship change if gender is considered?
    - Do students' whose parents are involved with Parent Portal have higher critical thinking scores than those who are not? (Data Source: Parent Services)
    - Do students who have engaged with the Career Center have higher critical thinking scores than those who have not? (Data Source: Career Center)
- Academic
    - Do students graduating from the Honors College have higher critical thinking scores than those who are not? (Data Source: Honors College)
        - If a differentiation is present, does it persist when controlling for demographic factors (e.g., ACT score)?
            - Technique required: Propensity Score Matching
    - Do students who attended a Library Information Session by an Instructional Librarian have higher critical thinking scores than those who did not? Does the frequency of engagement matter? (Data Source: Instructional Librarians)
    - Do students who are heavier users of Canvas have higher written communication outcomes (Data Source: Canvas)?
- Note – similar questions can be asked of First Destination Survey data. Likewise, the FDS data can be connected to this dataset (e.g., do students with higher critical thinking skills earn higher initial salaries?)