The University of Iowa

Iowa Testing Programs

College Readiness with the Iowa Assessments

ITP Research Series

Anthony D. Fina
Catherine J. Welch
Stephen B. Dunbar
Timothy N. Ansley

ITP Research Series

2015.1

Abstract

This study illustrates how the Iowa Assessments may be used to provide college readiness information on Iowa students. Compared to students who were not deemed college-ready, students who met the Iowa Benchmarks were up to twice as likely to earn letter grades equal to B or above in entry-level college courses in each subject area. Students who met the benchmark in all three subject areas had an average first-year grade point average of 3.28, or about a B+. In comparison, students who met at least one of the benchmarks earned a B average their first year, suggesting that the Iowa benchmarks provide information on readiness in a specific domain and overall college readiness in their first year as well.

College Readiness with the Iowa Assessments

Recent initiatives, such as the Common Core State Standards, Race to the Top, and several state-level reforms, have placed an emphasis on college and career readiness (e.g., U.S. Department of Education, 2010). They have been designed to make sure that students graduating from high school have the necessary skills for whatever path they may choose after graduation. These reforms recognize that simply earning a high school diploma does not necessarily mean that a student is ready for college-level work (National Conference of State Legislatures, 2014; Rado, 2014; Sheehy, 2012). For example, less than 40% of 12th grade students nationally have the required mathematical and reading skills needed for entry-level college courses, according to the National Assessment of Educational Progress (Driscoll, 2014).

To address these shortcomings, many states have examined how their state assessments might be used to improve the college readiness of students. In all, 48 states have adopted more rigorous college and career readiness standards, and 10 have also examined how their state-developed assessments can assess college readiness (Education Commission of the States, 2014). In Iowa, the relationship between scores on its state tests, referred to here as the Iowas (IAs), and performance in college has long been investigated (e.g., Fina, 2014; Furgol, Fina, & Welch, 2011; Qualls & Ansley, 1995; Scannell, 1960).

Past research studies have used a variety of statistical methods to relate test results to success in college. The most common measures of success in college are course grades and grade point average (GPA). In general, assessment results have a positive relationship with performance in college, be it college course grades or first-year GPA (FYGPA), even after statistically controlling for other factors. This indicates that the higher a student's test score is,

the more likely the student possesses the knowledge and skills required to be successful in the first year of college.

Given this positive relationship, it is important that students, parents, and school personnel have access to test score information as it relates to college success. Benchmarks can be a useful way to achieve that goal. In particular, benchmarks can quickly convey a student's readiness standing relative to some likelihood of future success. For benchmarks to be relevant however, they need be defined using an appropriate sample and criterion. Thus, the purpose of this study was to set empirical benchmarks relevant to Iowa students that signify the level of performance needed on the IAs to be successful in entry-level, credit-bearing college courses.

Benchmark Study

Data Sources

Data for this study came from two primary sources. The first source was the records of Iowa Testing Programs (ITP). These contained scores on the IAs from students who graduated from high school in the spring of 2009. The second source of data was the Office of the Registrar at the University of Iowa. This included records for all first-time, first-year students who enrolled at the university in the fall of 2009. The dataset contained course grades and FYGPA. Datasets were merged using first and last name, birth month, birth year, and sex. There were 1,534 students who were identified as having a Grade 11 National Standard Score (NSS) in Reading and a grade in a corresponding general education course. Likewise, there were 1,421 students in the Math dataset with valid data on both variables, and 1,279 students in Science.

The courses selected for setting the benchmarks were part of the General Education (GenEd) program at the University of Iowa. All students at the University of Iowa must complete these GenEd requirements to be eligible for graduation. For a list of the courses used in setting

the benchmarks, see Appendix A. Courses that would normally be identified as Social Science courses were used to set the Reading benchmark because these courses typically have a heavy reading load. The most common course was Elementary Psychology. In Math where initial course placement depends on students' prior work, higher-level courses that required a GenEd course as a prerequisite were also included because enrollment in these courses would fulfill the GenEd requirements in Quantitative and Formal Reasoning. The most common courses were Math for Business and Math for the Biological Sciences. In Science, the most common courses were General Chemistry and Principles of Biology.

Research Analyses

In this study, the relationships between grades and scores on the IAs were examined first. Next, grades in typical credit-bearing GenEd courses were modeled as a function of test scores. To model this relationship, course grades were divided into two categories: "success" and "failure." Then a logistic regression model was used to determine the probability of success in typical first-year courses at the University of Iowa for a given test score. For each subject area, a cutoff score was selected such that the probability of a grade of B or higher in the course was .50. Success was defined as a .50 probability because it is common today for more than 50% of students to earn grades of A or B, and thus reflects observed grading practices.

Results

To examine the relationship between course grades and test scores, graphs were made that displayed the average National Standard Score (NSS) by earned letter grade. These are provided in Figure 1. It is apparent that higher course grades are associated with higher average test scores. In Reading and Science, the progressively higher averages moving up the grade scale suggest a stronger relationship between grades and test scores. In Math, the bars may increase

more gradually because of the larger range of courses used to set the benchmarks, as described earlier. The highest averages associated with all letter grades were in Science. A's were associated with an average NSS value of 348, B's with 337, and C's with 325. Results in Math and Reading were similar to each other and generally 10-15 points lower than Science. Across subject areas, the majority of students earned a C or better (about 82%). Only about 50% of students earned a B or higher in Reading and Math (44% in Science).

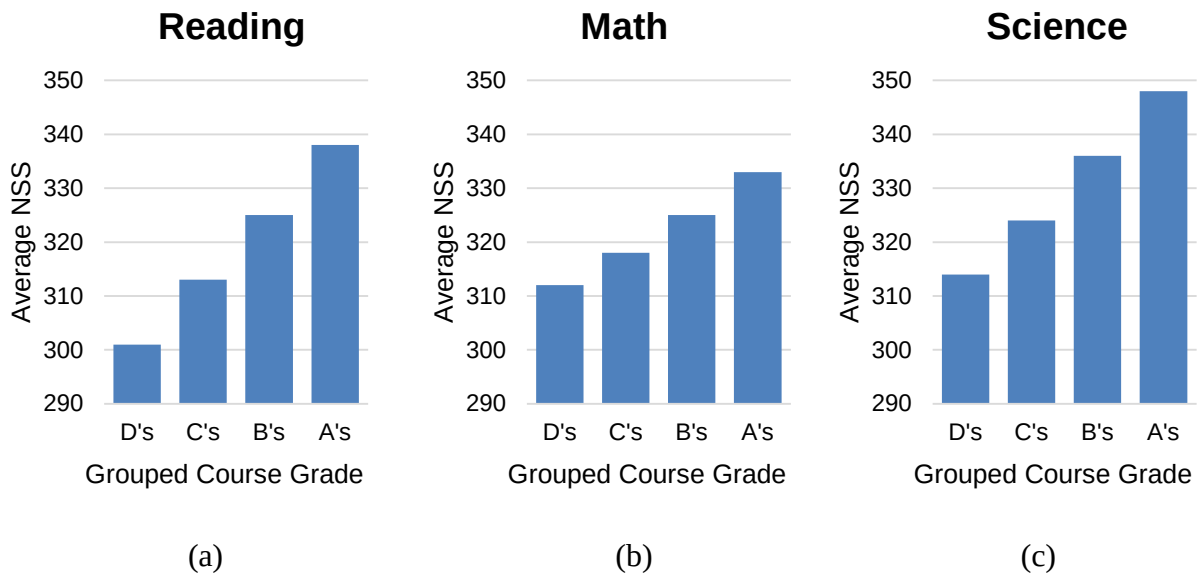

Figure 1.
Average NSS on the IAs by course grade.

Next, logistic regression was used to set the benchmarks. Logistic regression models the log odds (i.e., the log of a ratio of two probabilities) to describe the relationship between the independent variable (test scores) and dependent categorical variable (dichotomized course grades). Conceptually, the logistic model is used to estimate the probability of being successful for every test score. Success was defined in this study as earning a B or above in the GenEd courses at the University of Iowa.

The benchmarks were determined by identifying the IA score that was associated with a .50 probability of earning a B or above[1]. These benchmarks are referred to here as the Iowa Benchmarks. The cut scores represent a target for college readiness that defines the level of achievement a student needs in order to enroll and succeed in typical entry-level credit-bearing courses in college. The resulting Iowa Benchmarks for Reading, Math, and Science were 317, 319, and 336, respectively.

To determine the extent to which the benchmarks separated students into a "college-ready" group and "not college-ready" group, Table 1 presents the performance of students who met the benchmarks and those who did not for comparison purposes[2]. In Reading and Math, 61% of the students who were classified as college-ready earned a grade of B or better, and about 88% of students earned a C or better. Of the students who scored above the benchmark in Science, about 57% earned a grade of B or better, and 90% earned a grade of C or better.

Table 1
Percent of Students at/above Certain Letter Grades in Each Subject Area

		Did Not Meet Benchmark		Met Benchmark	
Benchmarks	Subject	C or Above	B or Above	C or Above	B or Above
Iowa	Reading	73	35	89	61
	Math	75	37	87	61
	Science	74	30	90	57
ACT[2]	Reading	65	25	85	53
	Math	70	31	86	58
	Science	71	28	91	57

The results in Table 1 show that the Iowa Benchmarks are useful predictors of success in first-year college courses. Note that IA scores above the cutoffs do not guarantee success (defined as B or higher) in the GenEd courses. Only about 60% of students who scored above the

[1] The results of these analyses are presented in Appendix B. In addition, the likelihood of getting a C or above is graphed as well.
[2]The ACT benchmarks are discussed in a later section, but are included in Table 1 to conserve space.

benchmarks earned a B or higher, though 89% did earn a C or higher. However, the usefulness of the benchmarks is supported by the fact that students who met the benchmark in a specific subject area were up to twice as likely to earn a B as students who did not meet the benchmark.

Next, the relationship between Iowa Benchmark attainment and performance in the first year of college was examined further. It was expected that students deemed college-ready in all three subject areas would have the highest average FYGPA, and this average would decrease as the number of benchmarks attained decreased. Students who were not deemed college-ready in any subject area were expected to have the lowest average FYGPA. Note that FYGPA only provides limited information on the ability of the Iowa Benchmarks to provide college readiness information because the Iowa Benchmarks represent readiness in domain specific areas and FYGPA is a broad measure of ability that reflects both cognitive and noncognitive factors.

The expected relationships were observed and the attainment of multiple benchmarks corresponded to higher FYGPAs (see Figure 2). Students who did not meet any Iowa Benchmarks, about 21% of the sample, had the lowest average FYGPA. For students who met all three benchmarks (35% of the sample), the average FYGPA was approximately equal to a B+. There were 21% of students who met one benchmark, and 23% who met two benchmarks. Interestingly, among students who attained only one benchmark, the average FYGPA across all three subject areas was 2.91, which is almost the equivalent of a B average. In addition, students who did not meet any benchmarks still had an average FYGPA (2.74) that many would find "reasonable." This is possible because students select courses tailored to their interests and abilities. A student might not have been successful in a specific GenEd course, but he or she may have performed better in other GenEd courses or electives that resulted in a higher, and "reasonable," FYGPA. That is, the Iowa Benchmarks are particularly useful for identifying

deficiencies in specific academic areas. It is also important to note that admissions decisions are based on a variety of factors. No single predictor can serve effectively in isolation.

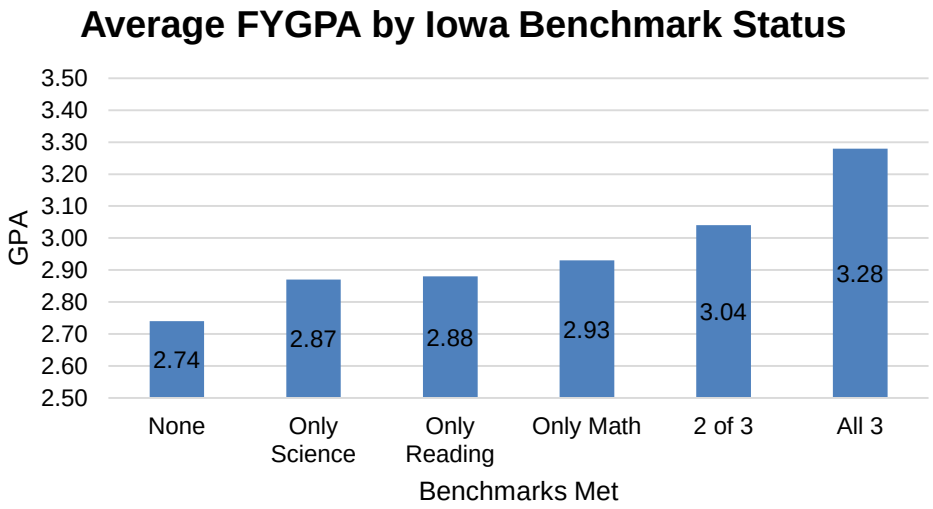

Figure 2
Average first-year GPA by number of benchmarks attained.

Comparison to the ACT College Readiness Benchmarks

The ACT Assessment also developed college readiness benchmarks (Allen & Sconing, 2005). These were investigated with the current sample for comparison purposes. The benchmarks were set using data from at least 14 four-year institutions and 17 two-year institutions. Compared to the sample used to set the Iowa Benchmarks, this sample was lower achieving based on median ACT scores in each subject area. The ACT benchmarks were 21 in Reading, 22 in Math, and 24 in Science. These benchmarks were used instead of the updated benchmarks (Allen, 2013) because these are the relevant benchmarks for when this cohort of students entered college (i.e., Fall of 2009). The results from using the ACT benchmarks are compared with results from using the Iowa Benchmarks described earlier.

The percentage of students earning a particular grade varied more for the ACT benchmarks than it did for the Iowa Benchmarks (see Table 1). While the percentages of students

earning a particular grade among students who met ACT's benchmarks were lower compared to the Iowa Benchmarks, the opposite was true for the students who did not meet ACT's benchmarks. This is a direct result of the Iowa Benchmarks representing a different criterion and is discussed next.

The importance of this distinction should not be understated. For example, unlike the Iowa Benchmarks, the expected relationships between FYGPA and benchmark attainment was not strictly observed (see Figure 3). Surprisingly, students who did not meet any of ACT's benchmarks, about 20% of the sample, had a higher FYGPA than students who met one benchmark (15% of the sample). This is likely a byproduct of ACT's benchmarks defining college readiness in terms of the skills needed "for typical students at typical colleges" to be successful (Allen & Sconing, 2005, p. 3), whereas the Iowa Benchmarks define college readiness in terms of the skills needed by Iowa students to be successful at the University of Iowa. This is possibly why more students attainted all three of ACT's benchmarks (45%) compared to the Iowa Benchmarks (35%).

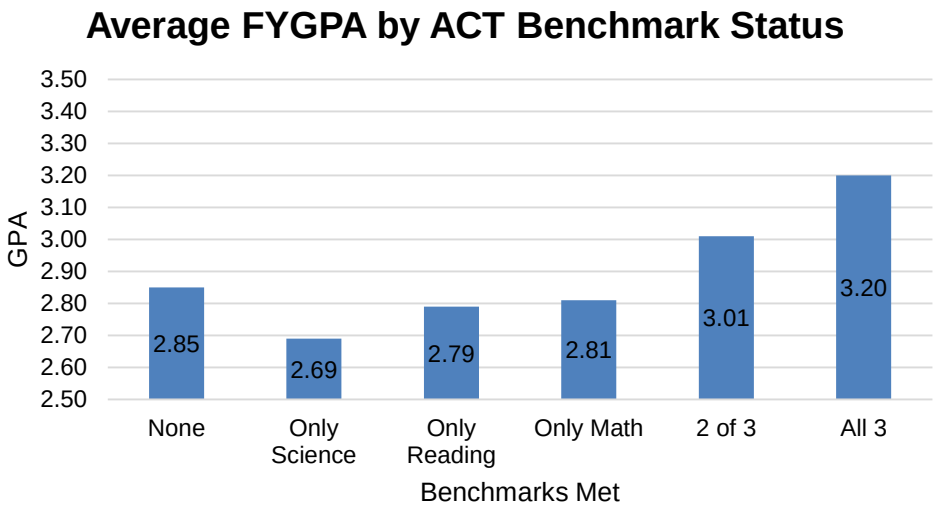

Figure 3
Average first-year GPA by number of ACT benchmarks attained.

Conclusion

Receiving a high school diploma does not guarantee that students have the necessary knowledge and skills required to be successful in college. To be successful in college requires a variety of skills and academic behaviors, in addition to content knowledge (see Conley, 2011). While noncognitive factors such as study skills and commitment are important for success in college, the academic knowledge and skills of students entering college are key to their eventual success. To help answer questions about how a student's level of achievement in broad domains in high school is related to performance in college, benchmarks were set on the IAs that defined the level of achievement a student needs to enroll and succeed in GenEd courses.

These benchmarks indicate the level of achievement in Grade 11 that is associated with a .50 probability of success in the corresponding GenEd courses taken at the University of Iowa. It was shown that the majority of students who scored above the Iowa benchmarks earn a B or higher and nearly all students earn at least a C across the three subject areas. While many students who did not meet a benchmark still passed the course, only about a third earn a grade of B or above. In addition, as the number of benchmarks met increases, so does average FYGPA. This study has shown how the IAs may be used for providing college readiness interpretations, such as in what areas students have the required skills to be successful in college and what areas require additional work. In addition, it has highlighted how there must be an alignment between the purpose of the benchmarks, the sample, and the criterion if the information provided by the benchmarks is going to be relevant to students and other stakeholders.

References

Allen, J., & Sconing, J. (2005). Using ACT Assessment scores to set benchmarks for college readiness. (ACT Research Report No. 2005-3). Iowa City, IA: ACT, Inc.

Allen, J. (2013). Updating the ACT college readiness benchmarks. (ACT Research Report No. 2013-6). Iowa City, IA: ACT, Inc.

Conley, D.T. (2011). *Redefining college readiness* (vol. 5). Eugene, OR: Educational Policy Improvement Center.

Driscoll, D. (2014, June 14). We must better prepare students for college. *Education Week.* Retrieved from http://www.edweek.org/ew/articles/2014/06/24/36driscoll.h33.html

Education Commission of the States. (2014). Blueprint for College readiness: A 50-state policy analysis. Retrieved from http://www.ecs.org/docs/BlueprintforCollegeReadiness.pdf

Fina, A.D. (2014). *Growth and the college readiness of Iowa students: A longitudinal study linking growth to college outcomes.* (Doctoral thesis). Iowa City, IA: University of Iowa.

Furgol, K., Fina, A.D., & Welch, C. (2011). *Establishing validity evidence to assess college readiness through a vertical scale.* Paper presented at the American Educational Research Association, New Orleans, LA.

Maruyama, G. (2012). Assessing college readiness: Should we be satisfied with ACT or other threshold scores? *Educational Researcher, 41*(7), 252-261.

National Conference of State Legislatures. (2014). Improving high schools through rigor, relevance, and relationships. Retrieved from http://www.ncsl.org/research/education/dropout-prevention-dropout-reporting-high-school.aspx

Rado, D. (2014, October 14). Most public school students aren't prepared for college work, data show. *Chicago Tribune*. Retrieved from http://www.chicagotribune.com/

Qualls, A.L., & Ansley, T.N. (1995). The predictive relationship of ITBS and ITED to measures of academic success. *Educational and Psychological Measurement, 55*(3), 485-498.

Scannell, D.P. (1960). Prediction of college success from elementary and secondary performance. *Journal of Educational Psychology, 51*(3), 130-134.

Sheehy, K. (2012, August 22). High school students not prepared for college, career. *U.S. News & World Report*. Retrieved from http://www.usnews.com/

U.S. Department of Education. (2010). *Race to the Top Assessment program: Executive summary*. Washington, DC: U.S. Department of Education.

Appendix A

Table A1
Courses Used to set the Benchmarks

Subject	Dept.	Course Number	Course Name	Benchmark Datasets	
				Percent	Count
Reading	06E	001	PRIN OF MICROECON	14.3	220
	06E	002	PRIN OF MACROECON	3.8	58
	019	090	SOC SCI FOUNDTN COM	3.5	53
	030	001	INTRO AM POLITICS	11.0	168
	030	045	INTRO COMP POLITICS	0.6	9
	030	050	INTRO TO POLIT BEHV	1.8	28
	030	060	INTR INTRNAT RELATN	2.6	40
	031	001	ELEM PSYCHOLOGY	32.5	499
	031	016	INTRO COGN PSYCH	0.7	10
	034	001	INTRO SOCIOL PRINC	13.6	208
	034	002	SOCIAL PROBLEMS	4.0	62
	034	020	PRIN OF SOC PSYCHOL	1.0	15
	036	070	COM EVERYDAY LIFE	1.0	16
	036	074	MEDIA AND SOCIETY	0.9	14
	103	011	LANGUAGE & SOCIETY	1.5	23
	113	003	INTRO STU CULT SOC	4.2	64
	113	014	LANG, CULT, & COMM	0.8	13
	169	070	PERSPEC LEISUR/PLAY	2.2	34
Math	22C	001	COMPUTER LITERACY	1.1	16
	22C	005	INTRO TO COMP SCI	5.0	71
	22C	016	COMPUTER SCIENCE I	0.5	7
	22M	005	TRIGONOMETRY	1.6	23
	22M	006	LOGIC OF ARITHMETIC	3.7	53
	22M	009	ELEM FUNCTIONS	2.7	38
	22M	013	MATH FOR BUSINESS	16.7	238
	22M	015	MATH BIOLOGICAL SCI	14.6	208
	22M	016	CALCULUS BIOLOG SCI	7.9	112
	22M	017	CALC MATRIX ALG BUS	8.1	115
	22M	025	CALCULUS I	6.9	98
	22M	027	INTRO LINEAR ALG	0.8	11

(Continued)

Subject	Dept.	Course Number	Course Name	Benchmark Datasets	
				Percent	Count
	22M	031	ENG MATH I: SV CALC	8.3	118
	22S	002	STATISTCS & SOCIETY	5.1	73
	22S	008	STAT FOR BUSINESS	3.9	55
	22S	025	ELEM STAT & INFER	8.3	117
	22S	102	INTRO TO STAT METH	0.6	8
	026	036	PRINC OF REASONING	4.2	60
Science	002	002	INTR ANIMAL BIOLOGY	6.6	84
	002	010	PRIN BIOLOGY I	15.2	194
	002	021	HUMAN BIOLOGY	9.3	119
	002	022	ECOLOGY & EVOLUTION	0.8	10
	004	005	TECHNOL & SOCIETY	3.0	39
	004	007	GENERAL CHEM I	31.0	397
	012	003	EARTH HIS & RESOUR	5.9	75
	012	007	AGE OF DINOSAURS	7.0	89
	012	008	INTRO ENVIRON SCI	7.8	100
	029	052	EXPLOR SOL SYS	2.8	36
	029	061	GEN ASTRONOMY I	0.9	11
	113	013	HUMAN ORIGINS	9.8	125

Appendix B

Figures B1-B3 plot the likelihood of getting a B or above in each subject area (the black line). In addition, the probability of getting a C or above is also plotted in each figure for comparison purposes. These likelihoods are more informative to students and other stakeholders compared to the dichotomous decisions offered by cut scores (Maruyama, 2012). Similarly shaped curves were seen across all three subject areas. At the low end of the scale, students are more than twice as likely to get a C in a course as they are to get a B. At the upper end of the distribution, students were generally likely to pass a course (over .95 probability of earning a C or above). However, the probability of those same students earning a B or above was only .80, approximately. As there are many factors influencing whether a student passes a course or not, it is not surprising that the probability of success never equals 1.0.

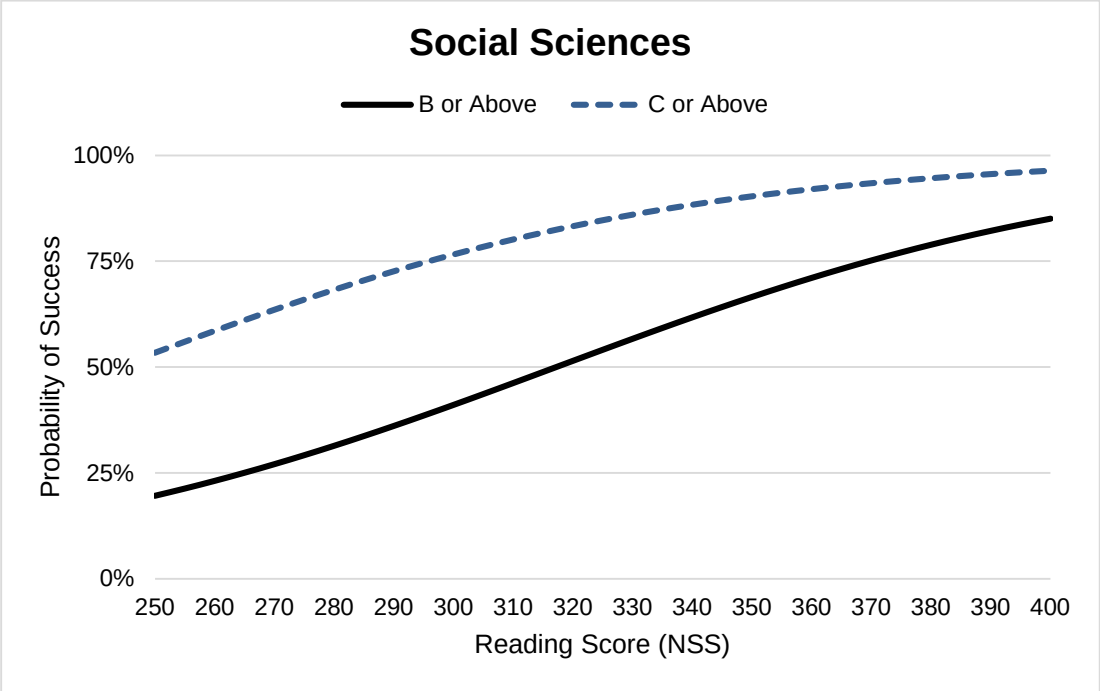

Figure B1
IA Reading score and probability of success in Social Science courses.

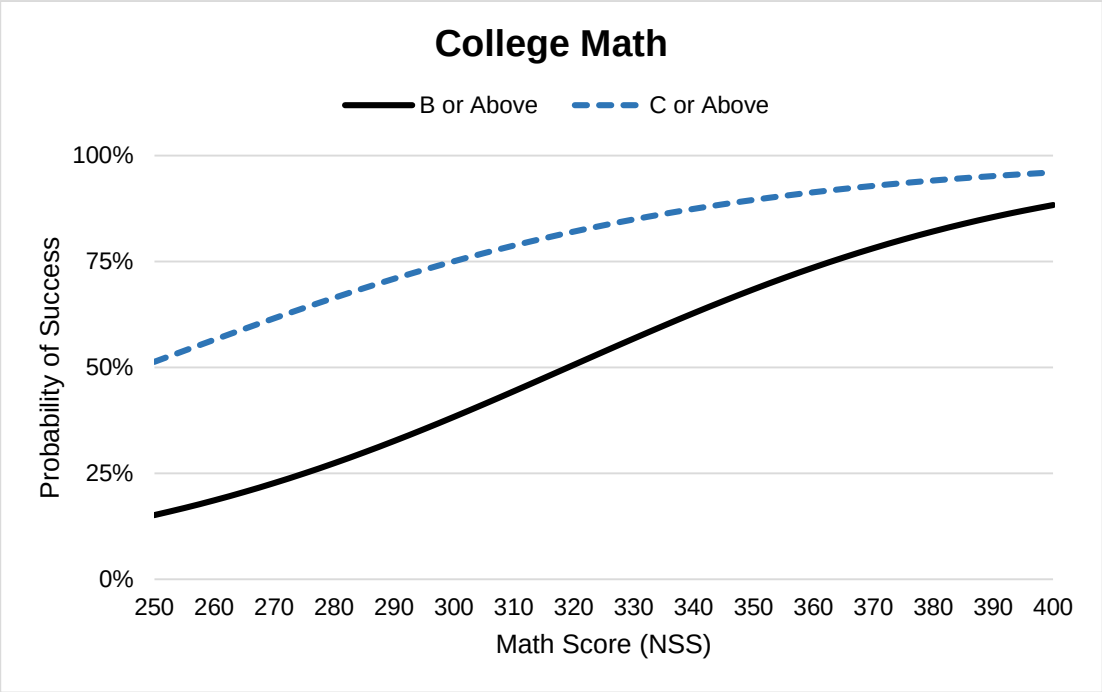

Figure B2
IA Math score and probability of success in College Math courses.

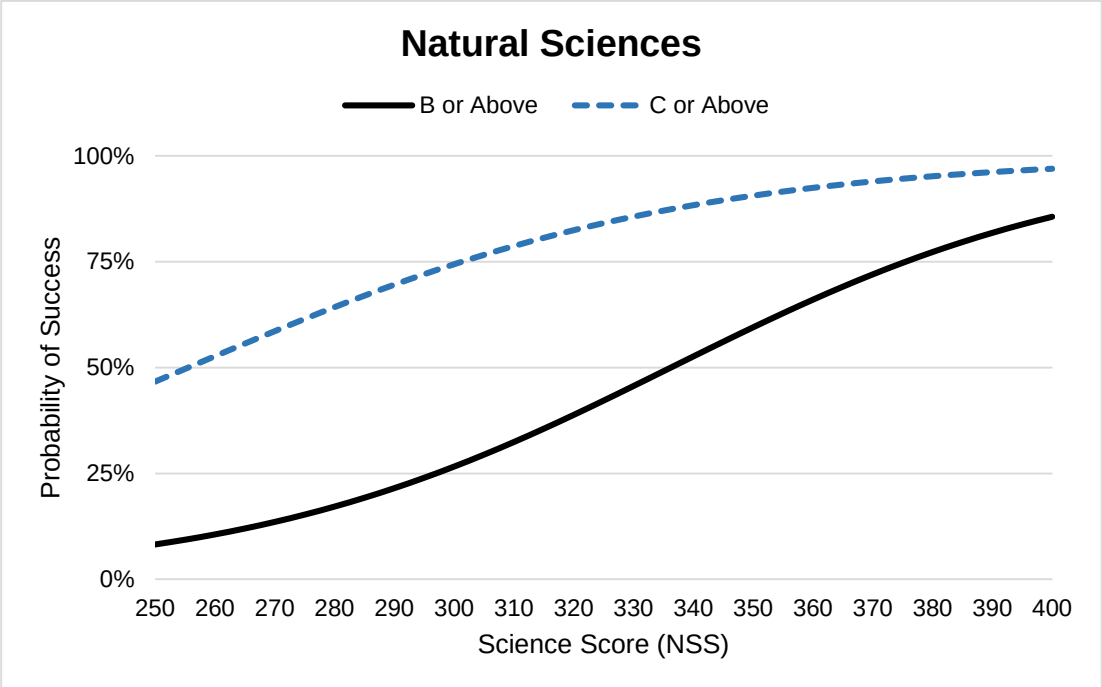

Figure B3
IA Science score and probability of success in Natural Science courses.