Exploring the Link Between Prerequisites and Performance in Advanced Data Structures

Sophia Krause-Levy, Sander Valstar, Leo Porter, and William G. Griswold
University of California, San Diego
{skrausel,avalstar,leporter,wgg}@eng.ucsd.edu

ABSTRACT

Recent work has identified a mismatch between instructor expectations of students' mastery of prerequisite course content and their actual ability. This invites the question of why this mismatch exists. We first examined grades in prerequisite courses and found they meaningfully correlated with performance on an assessment testing their knowledge of prerequisite material. In addition, we found neither taking alternatives to the primary identified prerequisites nor the delay between taking prerequisite courses and the follow-on course meaningfully impacts performance. Second, we confirmed that prerequisite course grades are significantly correlated with the grade in the follow-on course—confirming that the grades in the previous courses convey some information about student understanding of those topics. Perhaps surprisingly, we found that grades in courses outside computing were similarly correlated as those courses inside computing, suggesting that underlying factors such as general study skills may be as important as the domain-specific knowledge itself.

CCS CONCEPTS

• **Social and professional topics → Computing Education**.

KEYWORDS

prerequisites, study habits, curriculum, knowledge retention

ACM Reference Format:
Sophia Krause-Levy, Sander Valstar, Leo Porter, and William G. Griswold. 2020. Exploring the Link Between Prerequisites and Performance in Advanced Data Structures. In *The 51st ACM Technical Symposium on Computer Science Education (SIGCSE '20), March 11–14, 2020, Portland, OR, USA*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3328778.3366867

1 INTRODUCTION

The role of prerequisite knowledge in student success has just begun to be explored in computer science (CS). A recent study conducted by our research group at our university, a large US research university, used pre- and post-term tests of prerequisite knowledge in an upper-division data structures course to find that the degree of preparation varies widely among students and that the gap in preparation translated to students performing worse [23].

Concerningly, our study also found a surprisingly low level of preparedness among many of the students, despite grades in prerequisite courses being relatively high. In particular, 29% of students "failed" the pre-course test, despite fewer than 1% of their prerequisite course grades being a D. This disconnect raised the question of how such low preparedness could be possible.

Our impression as instructors and teaching assistants is that prerequisite preparation varies widely. On the assumption that knowledge fades with time [3], we worried that the length of time between taking a prerequisite and the follow-on course could be hurting some students. We also had a concern that because our CS program allows for multiple ways to satisfy some prerequisites—most not under our program's direct control—that some students might be disadvantaged in their prerequisite preparation.

In a different vein, it was also our impression that scores on exams tend to be lower than scores on other assessments. Collaboration on assignments could be contributing to higher scores for some, as could low-stakes assessments that measure effort more than knowledge. Moreover, we had the impression that "good" students possess durable qualities such as study skills, personal habits, and motivation that play a substantial role in improving performance [10]. Taken together, we hypothesized these qualities would contribute to not only knowledge acquisition, but also completing course elements with the highest possible grade.

Studying the same course as the previous study, we tested these hypotheses by collecting comprehensive course and grade information, and applying a variety of statistical analyses. First, we examined how prerequisite course grades correlated with the pre-term prerequisite test, final exam grade, and course grade. We then did the same analysis for grade aggregates such as prerequisite grade point average (GPA) and non-prerequisite GPA (excluding other CS courses). GPA represents the average grade a student has earned in their courses over time. To gain an understanding of the relative contributions of the various inputs, we then examined the relative importance of the precursors via multiple linear regression.

We come away with two major conclusions. *Students Learn in Prerequisite Courses*: Although grades may overstate how much has been learned, we find that prerequisite grades are predictive of prerequisite knowledge on entry to the course (0.48 Pearson's correlation). Related to this, we find no deleterious effects on prerequisite knowledge due to the passage of time or taking alternative prerequisites. *Good Students are More Skilled at Taking Courses*: Prerequisite grades are even more predictive of the follow-on course grade (0.53 correlation), despite the course grade including knowledge of new content learned during the course. Somewhat surprisingly, we also find that a student's non-prerequisite GPA, excluding other CS courses, is also highly predictive of the course grade (0.51). Combined with additional statistical tests, this supports our hypothesis

SIGCSE '20, March 11–14, 2020, Portland, OR, USA

ACM ISBN 978-1-4503-6793-6/20/03…$15.00
https://doi.org/10.1145/3328778.3366867

that course grades reflect not just subject matter knowledge, but also underlying abilities to succeed on course assessments.

We next discuss previous work. The subsequent two sections describe our study design and results, followed by sections on the study's limitations and the implications of our work.

2 PREVIOUS WORK

2.1 Role of Prerequisite Courses

Prerequisite courses are deeply ingrained in higher education, resulting in a number of studies assessing their impact. Several have involved examining how students who have and have not taken prerequisite courses perform in subsequent courses [5, 9, 16–18, 20, 29]. Others have explored the value of remedial courses in preparing students for subsequent courses [1] and examining the relationship between undergraduate prerequisite course grades and performance in a graduate course in a business program [2, 6, 13].

Similar to our investigation, prior work outside CS has examined the relationship between course grades in prerequisites and follow-on courses, finding that prerequisite grades predict subsequent course grades [7, 12, 13, 24]. Wright et al. found that biochemistry course grades are not impacted by taking the recommended organic chemistry prerequisite, but did find a strong relationship between cumulative GPA and biochemistry course grade [29]. Similarly, we found a relationship between non-prerequisite non-major courses and subsequent course grade.

Recent work in biology by Shaffer et al. classified material according to the depth of coverage in prerequisite courses and then asked questions on this knowledge in subsequent courses. Compared to students who did not take a prerequisite, students who took the prerequisite showed superior performance on only the most deeply covered material [19]. Sato et al. used this method, combined with a quantitative and qualitative analysis of student feedback, to determine the value of taking a recommended microbiology lecture class before a microbiology lab [18].

In CS, prerequisites can play many roles, including ensuring students have the proper background knowledge or sufficient maturity, that they display a certain degree of commitment to a program, or to convey program requirements [25]. In particular, the role of mathematics and introductory computing prerequisites and corequisites have been well discussed [16, 25]. CS course prerequisites are described in the ACM curricula recommendations [8].

2.2 Examining Prerequisite Knowledge in CS

In previous work, we studied the prerequisite knowledge of 208 students in an upper-division data structures course through pre- and post-term surveys (tests), comparing them to final exam scores [23]. We found prerequisite proficiency on entry to the course was far lower than anticipated. Nearly a third of students (29%) demonstrated low proficiency and only 27% demonstrated high proficiency. However, students improved their proficiency over the term by 8 percentage points on average, with over a quarter of the low-proficiency students improving to at least medium proficiency by the time of the final exam. With these survey results, we found that both incoming prerequisite proficiency and improvement in prerequisite proficiency over the term are correlated with final exam performance.

2.3 Why Students May Not Learn Prerequisites

Many factors may cause lower than expected incoming prerequisite knowledge. One such possible cause is long knowledge retention intervals, well studied in the education literature (see [3]). Evidence points to large knowledge losses even during a short retention interval. After an extensive review of literature on knowledge retention, Custers concluded that although different studies report vastly different results, there is support to the claim that about a third of knowledge is lost during the first year, half is lost after a few years, and loss levels off after that [3]. Custers also noted there is evidence for much worse retention for useless knowledge, citing Ebbinghaus's forgetting curve, which shows a knowledge loss of almost 80% after one month [3, 4].

A study on retention of CS1 core concepts at the start of CS2 found that students who take CS1 in fall and CS2 in spring have a knowledge loss of 4%, whereas students taking CS1 in spring and CS2 in fall have a loss of 15.1%. This difference can be explained by the summer vacation between the spring and fall terms [22].

2.4 Success Factors Beyond Prerequisites

There is a long line of work examining myriad other factors that influence student success in introductory programming courses, including prior knowledge, gender, motivation, self-efficacy, and achievement goals [11, 15, 28, 30]. Although each of these factors have been shown to be influential, Watson et al. recently cautioned that these relationships may not always replicate [26], which has led to large replication efforts [31].

3 STUDY DESIGN

3.1 Hypotheses

We formalize the hypotheses stated in the Introduction as follows:[1]

H-PL: (P)rerequisite (L)earning. Students learn prerequisite subject matter in the prerequisite courses for a CS course.

H-PT: (P)rerequisite (T)ime. The length of time between taking a course's prerequisites and taking the course itself negatively affects student success in that course.

H-AP: (A)lternative (P)rerequisites. Satisfying a course's prerequisite with an alternative to the primary identified course negatively affects student success in that course.

H-OF: (O)ther (F)actors. Other factors, unrelated to subject matter knowledge, strongly affect a student's success in that course.

Although H-PL seems obvious, we include it to establish a backdrop for analyzing the other hypotheses. For example, it would be difficult to establish the relative importance of factors unrelated to prerequisite knowledge (H-OF) without understanding the contribution of learning prerequisite knowledge in prerequisite courses.

3.2 Course Context

Previously we investigated an upper-division Advanced Data Structures (ADS) course that itself is a prerequisite to many other upper-division courses. The course was taught at a large US research university (R1) that operates on the quarter system. Courses meet for 10

[1] In hypothesis testing it is standard to formulate a negation of the hypothesis, the null hypothesis, to be falsified. For example, the null hypothesis for H-PL is that students do not learn prerequisite knowledge or even unlearn it. In this paper, the null hypotheses are straightforward so we do not explicitly state them.

Table 1: Majors breakdown. Double majors are counted twice. Columns may not add up 100% due to rounding.

	All in ADS		Included in Study		Main-Track		Alt-Track	
Major	# Students	% Students	# Students	% Students	# Students	% Students	# Students	% Students
Math-Computer Science	133	32.13%	82	36.94%	31	24.80%	51	52.58%
Computer Science	109	26.33%	55	24.77%	44	35.20%	11	11.34%
Computer Engineering	39	9.42%	25	11.26%	22	17.60%	3	3.09%
Other	133	32.13%	60	27.03%	28	22.40%	32	32.99%
Total	363		202		120		82	

weeks, immediately followed by a week of final exams. Being an advanced course, it concerns the implementation of high-performance data structures such as hash tables, complex tree structures like B-Trees, and graphs. Course assignments were performed in C++, so the requirements for performance led to additional course material on the use of pointers and memory management. The course assessments were online reading assignments (10%), clicker participation (5%), in-class quizzes (5%), programming assignments (35%), midterm (20%), and final exam (25%). We observe that 50% of the course grade allows for collaboration or rewards effort. This proportion is not unusual in this CS program.

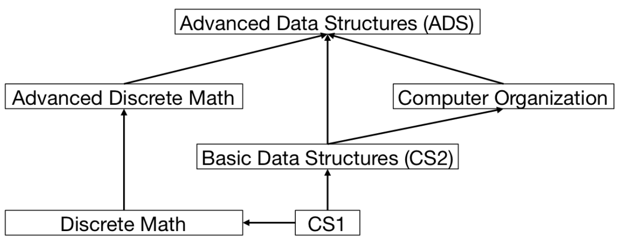

Figure 1: Prerequisite structure for primary identified prerequisite courses for Advanced Data Structures.

Figure 1 summarizes the lower-division primary prerequisite structure as it pertains to ADS. The course's main prerequisites are fairly traditional courses such as basic data structures (CS2) and advanced discrete math. CS1 is taught as a one or two course sequence, which the students are advised to select according to their prior preparation. CS1 is a co-requisite for Discrete Math. There are designated "alternative" courses satisfying most requirements (relevant for H-AP). Alternatives for Discrete Math include 'Introduction to Discrete Mathematics' and 'Mathematical Reasoning' in the Mathematics department. Alternative courses for Advanced Discrete Math include 'Discrete Mathematics and Graph Theory' and 'Enumerative Combinatorics' in the Mathematics department. Equivalent courses for Computer Organization include 'Introduction to Programming I' taught in C by the CS department, 'C/C++ Programming' in the Mechanical Engineering department and 'Engineering Computation' in the Electrical Engineering department.

3.3 Population

Although our population of students comes from the same ADS course previously studied by our research group, it contains a slightly different group of students than the cohort used in the previous analysis. To enable our previous analysis, our inclusion criteria included students who took the pre-term survey, post-term survey, and the final exam. Our current hypotheses do not require the post-term survey, but do require other data that not all students

Table 2: Non-prerequisite courses taken by ADS students.

Course ID	# Courses	% Courses
Math	965	33.47%
Writing	423	14.67%
Physics	200	6.94%
Cog. Sci.	190	6.59%
Economics	174	6.04%
Chemistry	131	4.54%
Other	800	27.75%
Average/Student	14.27	

possessed, resulting in a different cohort. Table 1 summarizes many characteristics discussed below.

Many of our research questions depend on prerequisite grades and the ability to compare the grades between different offerings of a given prerequisite. This led us to exclude students who took some prerequisite courses at a community college or different university, as we did not have access to the official equivalencies of those courses nor the grade distributions for those course offerings. We hope to gain access to this information for future work.

The pre-term survey and final exam were taken by 322 students. Of those, 206 took all of their prerequisites for ADS at our university. Of those, four took only one or two non-CS courses outside of their prerequisite courses for ADS. These students were excluded from the study population as our analysis relies heavily on a student's non-prerequisite non-CS GPA, and their GPAs would not provide enough information to be reliable. The remaining 202 students form the population for the analysis in this paper. Table 2 summarizes the kinds of non-prerequisite non-CS courses taken by these students.

To assess our hypothesis about the effects of taking alternatives courses to the primary identified prerequisite for ADS, we created two sub-populations. A student is *main-track* if they took all of the primary identified prerequisite courses for ADS (120 students). A student is *alternate-track* if they took all of the required prerequisites, but at least one of their prerequisites is a designated alternative to the primary identified prerequisite course (82 students).

3.4 Data Collection and Cleaning

Our study expands on the pre-term survey and final exam data collected for our previous study [23]. In addition, with help from our approved human subjects protocol, we obtained the students' grades in ADS as well as grades for all courses that the students had taken at our university prior to taking ADS.

We used the students' grades to calculate the average grade for the prerequisite courses required to take ADS (*prerequisite GPA*) and the average grade for all courses taken at our university excluding their prerequisite courses and any other CS courses taken at our university (*non-prerequisite GPA*). Because the purpose of the non-prerequisite GPA is to operationalize the skills of taking courses independent of prerequisite knowledge, CS courses are excluded from the non-prerequisite GPA to lessen contributions related to (a) special enthusiasm for CS or (b) prior CS knowledge that might implicitly overlap with the prerequisites. When calculating prerequisite GPAs, if a student took more than one course that counted for a single prerequisite, we chose the course that was taken most recently, as it best reflects prerequisite knowledge when entering the ADS course. This occurred 86 times (40% alternate-track, 60% main-track), for approximately 8.5% of the prerequisite courses. Neither GPA included courses that were taken pass/fail.

At our university grades are recorded on a 4 point scale where a 4.0 represents an A and 0.0 represents an F. Using grade points to represent knowledge or accomplishment is problematic as most students who receive an F did not earn a 0% in the instructor's gradebook. To account for this bias in grade point, we rescaled the 4.0 scale such that an F is equivalent to 50%.

All the data used in this study was normalized via z-score. For the courses outside ADS, the student grades we have are just a sample of the grades earned in those courses, not permitting a straightforward calculation of a z-score. Normalizing these is especially important because not all courses or course offerings are equally difficult. That is, an A in one course offering can mean something different than an A in another course offering. Fortunately, we were able to scrape our university's online course evaluation system for the grade distributions of all relevant courses and use them to compute the z-score for each grade (A, A-, B+, B, B-, C+, C, C-, D, F). Because we only have access to the course and term in which a student took a course, in cases where multiple sections of a course were offered the same term, we used the average of the distributions for those sections. In the cases where a distribution was not available at all, we used the average distribution for terms for which we had data. This occurred in just 10 sections across 3 different courses. In total there were 1223 sections across 477 courses.

4 RESULTS

4.1 H-PL: Students Learn in Prerequisites

To understand how prerequisites contribute to learning prerequisite subject matter, we calculated the Pearson correlation between prerequisite GPA and pre-term survey score, which is significant with a correlation of 0.48 (Table 3, under the heading "All"). As expected, it is more highly correlated than the other pre-term measure, non-prerequisite GPA (0.35). To assess whether these correlations are meaningfully different, we applied Williams's Test, which is used to test whether two correlations for the overlapping (i.e., dependent) populations are in fact different [27]. As shown in Table 4, the differences in the correlations are significant.[2] Together, these support the hypothesis that performance in prerequisite courses positively affects learning of prerequisite knowledge in CS.

[2]Williams's Test uses only the significance test, not the test value.

Table 3: Pearson correlations between precursor course measures and pre-term survey score. A star (*) indicates a significant correlation for $p = 0.05$.

	All		Main-Track		Alt-Track	
GPA Variable	corr	p	corr	p	corr	p
Prereq	0.48	<0.01*	0.50	<0.01*	0.46	<0.01*
NonPrereq	0.35	<0.01*	0.34	<0.01*	0.35	<0.01*

Table 4: Williams's Test for whether the correlations of the two precursors to survey score (Table 3, under heading "All") are different. A star (*) indicates significance for $p = 0.05$.

Variable 1	Variable 2	stat	p
Prereq GPA	NonPrereq GPA	3.20	<0.01*

Table 5: Pearson correlation for performance in prerequisites minus survey score with time since courses taken. A star (*) indicates a significant correlation for $p = 0.05$.

Variable	corr	p
Pre-Term Survey Score - Main Prereq GPA	0.08	0.25

4.2 H-PT: Prerequisite Learning Fades

Now that we know students are learning in prerequisites, we consider whether the fading of knowledge over time could be contributing to lower performance on the pre-term survey. We operationalized the passage of time as the sum of the months elapsed since taking the main prerequisites, Advanced Discrete Math, CS2, and Computer Organization (Figure 1). We excluded the other two prerequisites on the presumption that their immediate successors refreshed the knowledge from those courses, given the tight dependence in subject matter. In our previous study we found ADS itself provided such a refreshing effect to prerequisite knowledge [23]. To operationalize the change in knowledge over time, we subtracted a student's GPA for the aforementioned main predecessors from the pre-term survey (both as z-scores). We then calculated the correlation between the variables for net time elapsed and change in knowledge (Table 5). It is not significant, and at 0.08, it is not in the hypothesized direction (as the time gap increases so does prerequisite knowledge), providing no support for the hypothesis.

4.3 H-AP: Taking Alternatives to Primary Prerequisites Diminishes Student Success

To determine whether there was a difference between the success of main-track and alternate-track students in ADS, we began by calculating, for each group, the correlation for every combination of precursor measure (prerequisite GPA, non-prerequisite GPA, and pre-term survey score) and the two result measures (ADS final exam and ADS grade). The correlations can be found on the right side of Table 6 under the headings Main-Track and Alt-Track. All of the correlations are significant, except for the main-track correlation between survey score and ADS grade.

Then, for each of the six combinations, we tested whether the correlation for alternate-track was significantly different than that for main-track. Because the test is between non-overlapping (i.e., independent) groups, we applied Fisher's r-to-z transformation to the two correlations and then performed Steiger's z-test [21] (Table 7). None of the correlation pairs are found to be different, as

Table 6: Pearson correlations for performance in course with a comparison to final exam score and ADS course grade: correlation with p-values. A star (*) indicates a significant correlation for $p = 0.05$.

	All				Main-Track				Alt-Track			
	ADS Final Exam		ADS Grade		ADS Final Exam		ADS Grade		ADS Final Exam		ADS Grade	
Variable	corr	p	corr	p	corr	p	corr	p	corr	p	corr	p
Prereq GPA	0.66	<0.01*	0.53	<0.01*	0.63	<0.01*	0.53	<0.01*	0.70	<0.01*	0.53	<0.01*
NonPrereq GPA	0.55	<0.01*	0.51	<0.01*	0.56	<0.01*	0.51	<0.01*	0.53	<0.01*	0.51	<0.01*
Pre-Term Survey Score	0.30	<0.01*	0.21	<0.01*	0.35	<0.01*	0.17	0.06	0.26	0.02*	0.27	0.02*

Table 7: Steigler's z-test for whether the main-track and alternate-track correlations are different ($p = 0.05$).

	ADS Final Exam		ADS Grade	
Variable	z-stat	p	z-stat	p
Prereq GPA	0.86	0.39	0.00	1.00
NonPrereq GPA	-0.29	0.77	0.00	1.00
Pre-Term Survey Score	-0.68	0.50	0.72	0.47

the test magnitudes are below 1.96 (and p > 0.05). As an example, the correlation for prerequisite GPA and ADS grade is 0.53 for main-track and 0.53 for alternate-track, with the z-test returning a result of 0.00 with p=1.00. Thus, we find no support for the hypothesis that alternative prerequisites negatively affect performance.

4.4 H-OF: Factors Other than Prerequisite Learning Boost Student Performance

If factors beyond actual subject matter knowledge contribute to student grades, we would expect to see a (strong) correlation between course grades in unrelated subjects. On the other hand, we would expect tests to more strongly correlate to subject matter knowledge since (a) collaboration is both prohibited and harder and (b) typically assess subject matter knowledge more than effort. Thus, to measure how other factors unrelated to subject matter knowledge affect student success in a course, we took the correlations between students' non-prerequisite GPA and three measures related to prerequisites: ADS course grade, ADS final exam score, and pre-term survey score (Table 6, under the heading "All"). All of the correlations are statistically significant.

If non-prerequisite GPA were nearly as predictive as prerequisite GPA for success in ADS, then factors beyond subject matter knowledge must be playing a major role in success. This could be validated if the correlations between ADS grade and the two GPAs were deemed similar. Indeed, Williams's Test finds no significant difference in the correlations (Table 8, first row, under the heading "ADS grade"). We would also expect Williams's Test to find significant differences involving exams, and this is indeed the case (Tables 4 and 8 second and third rows, heading "Final Exam").

Having established these relationships, the question of the precursors' relative importance in outcomes arises. To this end, we performed two multiple linear regression analyses, one for the dependent variable (DV) ADS grade and the other for ADS final exam, both with the independent variables (IVs) non-prerequisite GPA, prerequisite GPA, and pre-term survey score (Table 9).

Since the independent variables are no-doubt interdependent, their model coefficients alone cannot determine the relative contribution of the independent variables. Nor are the coefficient p-values considered useful, as they don't capture effect size. The relevance of an independent variable is determined instead by how it effects the model's R-squared value. Thus we employed the method of calculating the increase in R-squared that each variable produced when it was added to a model that already contained all the other independent variables (i.e., the one other variable in this case). This is called the *R-squared difference*.

Our expectations are borne out by the data, with prerequisite GPA dominating non-prerequisite GPA for ADS final exam (coefficient 1.10 vs. 0.15 and r^2 difference 0.1229 vs. 0.0021), but with much more parity for ADS grade (coefficient 0.60 vs. 0.41 and r^2 difference 0.0479 vs. 0.0193).

In conclusion, we accept the hypothesis that factors outside subject matter knowledge are playing a large role in student success. Separately, we observe that pre-term survey score is not useful in the model. We suspect that prerequisite GPA sufficiently captures prerequisite knowledge on entry to the course, since H-PL is supported and H-PT and H-AP are not.

5 LIMITATIONS AND RISKS TO VALIDITY

Our results are based on a single offering of a particular course at a single research university, limiting generalizability. However, the curriculum is quite common and informed by the ACM Curriculum guidelines [8]. Moreover, US research-intensive universities teach a large proportion of CS students on the continent, making this an interesting demographic in itself. Still, results could vary outside this context, encouraging attempts to replicate our results.

At a more specific level, our conclusions on the effects of taking alternative prerequisites only included on-campus alternatives. Variations in curricular match and instructional quality are possibly higher for courses from other campuses, limiting generalizability.

There are a few possible risks to the validity of our conclusions. For one, the pre-term survey is short and not validated. This could result in over- or under-estimating prerequisite knowledge, depending on the actual shortcomings of the survey. Also, beyond goodwill and pride, the students had no incentive to perform well on the survey, meaning that students lacking these motivations could have under-performed. To mitigate these, we did not use the magnitude of the survey score in drawing our conclusions, only the correlation, which depends on co-variance. Of lesser concern is that we could not include all students who took ADS in the study, as explained in Section 3.3. Perhaps students who come to class are more highly motivated than those who don't, but it could also be the case that high-performing students don't feel the need to come to class every day. The motivation element increases the chance that the students who took the survey would take it more seriously, mitigating the just-cited incentive risk. The additional factors for exclusion did not appear to significantly alter the population mix (Table 1).

Table 8: Williams's Test for whether precursors are differently correlated to outputs. A star (*) indicates significance for $p = 0.05$.

		ADS Final Exam		ADS Grade	
Variable 1	Variable 2	stat	p	stat	p
Prereq GPA	NonPrereq GPA	3.16	<0.01*	0.52	0.61
Prereq GPA	Pre-Term Survey Score	6.49	<0.01*	5.17	<0.01*
NonPrereqGPA	Pre-Term Survey Score	3.66	<0.01*	4.26	<0.01*

Table 9: Multiple linear regression statistics for dependent variables of final exam score and ADS course grade. A star (*) indicates a significant correlation for $p = 0.05$.

	DV: ADS Final Exam (r^2: 0.437)					DV: ADS Grade (r^2: 0.305)				
Independent Variable	Coefficient	Std. Error	t-stat	p	r^2 diff	Coefficient	Std. Error	t-stat	p	r^2 diff
Prereq GPA	1.10	0.17	6.57	<0.01*	0.1229	0.60	0.16	3.69	<0.01*	0.0479
NonPrereq GPA	0.15	0.18	0.86	0.39	0.0021	0.41	0.17	2.34	0.02*	0.0193
Pre-Term Survey Score	-0.01	0.06	-0.18	0.86	0.0001	-0.04	0.06	-0.72	0.47	0.0018

A risk in distinguishing the main-track and alternate-track groups is that the groups are biased. As can be seen in Table 1, there are substantially more CS majors in the main-track group. One possible source of bias is that entry into the CS major is competitive, so it is possible that the main-track group would perform better in CS courses than the alternate-track group. However, our analysis for H-AP depends on the correlation of grades to survey score, and the value of a correlation does not depend on the relative magnitude of grades, but rather co-variance of the magnitudes of grades. In other words, despite a possible bias in the populations, the alternate-track group was indistinguishable from on the main-track group in terms of how grades predict performance.

For main-track and alternate-track students, 8.5% of students took and passed two different courses that satisfied a prerequisite. Additionally, an unknown number of students may have taken a late drop in a course satisfying a prerequisite, or have taken a course satisfying a prerequisite at a community college. These could disguise shortcomings in the prerequisite courses that we used in our main-track/alternate-track prerequisite GPA calculations.

Our use of grade distributions to normalize grades across offerings, while controlling for the difficulty of a course offering, also potentially masks actual differences in student performance. However, the courses in our study have enough students (120 students on average) that a course offering likely represents a substantial, and representative, sample of the student body.

Finally, although we concluded that factors related to subject-matter knowledge are a large contributor to student success in a course, another possible cause is that students are acquiring ADS-relevant subject matter in non-prerequisites. We excluded CS from the non-prerequisite GPA to mitigate this risk, and it's unlikely that there would be a measurable pattern given the diverse non-prerequisite demographics we found in our population (Table 2). Still, this shadow-prerequisite question is deserving of future study, especially for non-prerequisite CS courses.

6 DISCUSSION AND CONCLUSION

This study began with two questions: how could students manifest so little prerequisite knowledge coming into a course, and how could they get such good grades with so little apparent preparation?

The good news from our study is that students are learning in prerequisite courses, and there is no evidence for the passage of time or taking alternative prerequisites diminishing that knowledge. Regarding the lack of ill effects of the passage of time, it should be kept in mind that the time between taking prerequisites and taking ADS is relatively bounded. For those with longer breaks, this knowledge might be refreshed by intervening activities such as internships, a job, or taking other CS courses.

Also interesting is that students who do better in non-prerequisite non-CS courses also do better in ADS. More specifically, students are distinguishing themselves not just through subject matter knowledge, but also other factors. Although this study couldn't reveal what those factors are, we expect influences from study habits, excessive collaboration, and low-stakes assessments (combined with student personal habits that maximize these assessments). Is this bad news? To the extent that it is giving students a grade bump that allows them to side-step learning valuable subject matter knowledge, yes. On the other hand, students who are excelling due to these other factors are demonstrating they know how to master a complex ecology to maximize a desired outcome (a grade). Those skills are likely transferable to other contexts, offering the promise of success after graduation. Future work should investigate these factors, their impact on knowledge acquisition, and why some students have these skills and habits and others do not.

Tantalizingly, students who do better in non-prerequisite non-CS courses also do better on the ADS final exam, although the effect is not as strong as on ADS course grade. We hypothesize two possibilities for this positive correlation. One is that non-prerequisite courses and ADS could be dependent on (and/or enhancing) "shadow prerequisites", for example spatial reasoning [14] or critical thinking skills, implying that courses outside of CS continue to improve CS skills. Another explanation is that exams, like other course elements, are often gameable, e.g., by smartly guessing which topics will get the most coverage on the exam and studying accordingly, or the use of test strategies such as process-of-elimination or matching the syntax of a question to a known practice question. These skills fit in our category of "other factors" for success, and could be quite transferable between course subjects, explaining the correlation. We leave the exploration of these hypotheses for future work.

ACKNOWLEDGMENTS

This work was supported in part by NSF IUSE Grant DUE-1712508.

REFERENCES

[1] Pansy W Aycaster. Factors impacting success in community college developmental mathematics courses and subsequent courses. *Community College Journal of Research & Practice*, 25(5-6):403–416, 2001.

[2] Askar Choudhury, Don Robinson, and Ramaswamy Radhakrishnan. Effect of prerequisite on introductory statistics performance. *Journal of Economics and Economic Education Research*, 8(3):19, 2007.

[3] Eugène J.F.M. Custers. Long-term retention of basic science knowledge: A review study. *Advances in Health Sciences Education*, 15(1):109–128, 2010.

[4] Hermann Ebbinghaus. *Über das Gedächtnis: Untersuchungen zur Experimentellen Psychologie (Memory: A Contribution to Experimental Psychology)*. Dunker and Humblot, 1885.

[5] Joseph P Forester, David L McWhorter, and Maria S Cole. The relationship between premedical coursework in gross anatomy and histology and medical school performance in gross anatomy and histology. *Clinical Anatomy: The Official Journal of the American Association of Clinical Anatomists and the British Association of Clinical Anatomists*, 15(2):160–164, 2002.

[6] Jeffrey J Green, Courtenay C Stone, Abera Zegeye, and Thomas A Charles. Changes in math prerequisites and student performance in business statistics: Do math prerequisites really matter? *Journal of Economics and Finance Education*, 6(2):27–38, 2007.

[7] Faridul Islam, Saleheen Khan, Ian Wilson, and Reed Gooch. The value of prerequisite courses for statistics. *The Journal of Business Inquiry*, 7(1):61–67, 2008.

[8] Association for Computing Machinery (ACM) Joint Task Force on Computing Curricula and IEEE Computer Society. *Computer Science Curricula 2013: Curriculum Guidelines for Undergraduate Degree Programs in Computer Science*. 2013.

[9] Colleen E Kauffman and David A Gilman. Are prerequisite courses necessary for success in advanced courses?. 2002.

[10] Soohyun Nam Liao, Sander Valstar, Kevin Thai, Christine Alvarado, Daniel Zingaro, William G. Griswold, and Leo Porter. Behaviors of higher and lower performing students in CS1. In *Proceedings of the 25th Conference on Innovation and Technology in Computer Science Education*, pages 196–202, 2019.

[11] Alex Lishinski, Aman Yadav, Jon Good, and Richard Enbody. Learning to program: Gender differences and interactive effects of students' motivation, goals, and self-efficacy on performance. In *Proceedings of the 2016 Conference on International Computing Education Research*, pages 211–220, 2016.

[12] Marshall A Martin. Course prerequisites and undergraduate student performance. *NACTA Journal*, 33(1):38–42, 1989.

[13] Amy McMillan-Capehart and Tope Adeyemi-Bello. Prerequisite coursework as a predictor of performance in a graduate management course. *Journal of College Teaching & Learning (TLC)*, 5(7), 2008.

[14] Jack Parkinson and Quintin Cutts. Investigating the relationship between spatial skills and computer science. In *Proceedings of the 2018 ACM Conference on International Computing Education Research*, pages 106–114, 2018.

[15] Vennila Ramalingam, Deborah LaBelle, and Susan Wiedenbeck. Self-efficacy and mental models in learning to program. In *Proceedings of the 9th Conference on Innovation and Technology in Computer Science Education*, pages 171–175, 2004.

[16] Christine F Reilly and Emmett Tomai. Should college algebra be a coerequisite for computer science 1? FECS: Position paper. In *Proceedings of the International Conference on Frontiers in Education: Computer Science and Computer Engineering (FECS)*, page 1, 2013.

[17] Allen A Rovick, Joel A Michael, Harold I Modell, David S Bruce, Barbara Horwitz, Thomas Adamson, Daniel R Richardson, Dee U Silverthorn, and Shirley A Whitescarver. How accurate are our assumptions about our students' background knowledge? *Advances in Physiology Education*, 276(6):S93, 1999.

[18] Brian K Sato, Amanda K Lee, Usman Alam, Jennifer V Dang, Samantha J Dacanay, Pedro Morgado, Giorgia Pirino, Jo Ellen Brunner, Leanne A Castillo, Valerie W Chan, et al. What's in a prerequisite? a mixed-methods approach to identifying the impact of a prerequisite course. *CBE—Life Sciences Education*, 16(1):ar16, 2017.

[19] Justin F Shaffer, Jennifer V Dang, Amanda K Lee, Samantha J Dacanay, Usman Alam, Hollie Y Wong, George J Richards, Pavan Kadandale, and Brian K Sato. A familiar (ity) problem: Assessing the impact of prerequisites and content familiarity on student learning. *PloS one*, 11(1):e0148051, 2016.

[20] Krista M Soria and Lori Mumpower. Critical building blocks: Mandatory prerequisite registration systems and student success. *NACADA Journal*, 32(1):30–42, 2012.

[21] James H. Steiger. Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87(2):245–251, 1980.

[22] Matthew Tennyson and Marc Beck. A study of knowledge retention in introductory programming courses. *Journal of Computing Sciences in Colleges*, 33(4): 13–20, 2018.

[23] Sander Valstar, William G. Griswold, and Leo Porter. The relationship between prerequisite proficiency and student performance in an upper-division computing course. In *Proceedings of the 50th Technical Symposium on Computer Science Education*, pages 794–800, 2019.

[24] J.D. Vitale, S.P. Wanger, and D.C. Adams. Explaining student performance in an undergraduate agricultural economics classroom. *NACTA Journal*, 54(1):2, 2010.

[25] Henry M. Walker. Prerequisites: Shaping the computing curriculum. *ACM Inroads*, 1(4):14–16, 2010.

[26] Christopher Watson, Frederick W.B. Li, and Jamie L. Godwin. No tests required: Comparing traditional and dynamic predictors of programming success. In *Proceedings of the 45th Technical Symposium on Computer Science Education*, pages 469–474, 2014.

[27] Eric J. Williams. The comparison of regression variables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 21(2):396–399, 1959.

[28] Brenda Cantwell Wilson and Sharon Shrock. Contributing to success in an introductory computer science course: A study of twelve factors. In *Proceedings of the 32nd Technical Symposium on Computer Science Education*, pages 184–188, 2001.

[29] Robin Wright, Sehoya Cotner, and Amy Winkel. Minimal impact of organic chemistry prerequisite on student performance in introductory biochemistry. *CBE—Life Sciences Education*, 8(1):44–54, 2009.

[30] Daniel Zingaro and Leo Porter. Impact of student achievement goals on cs1 outcomes. In *Proceedings of the 47th Technical Symposium on Computing Science Education*, pages 279–296, 2016.

[31] Daniel Zingaro, Michelle Craig, Leo Porter, Brett A. Becker, Yingjun Cao, Phill Conrad, Diana Cukierman, Arto Hellas, Dastyni Loksa, and Neena Thota. Achievement goals in CS1: Replication and extension. In *Proceedings of the 49th Technical Symposium on Computer Science Education*, pages 687–692, 2018.