JOURNAL OF CHEMICAL EDUCATION

pubs.acs.org/jchemeduc Article

Beneath the Surface: An Investigation of General Chemistry Students' Study Skills to Predict Course Outcomes

Emily L. Atieh, Darrin M. York, and Marc N. Muñiz*

Cite This: J. Chem. Educ. 2021, 98, 281−292 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: As the conversation in higher education shifts from diversity to inclusion, the attrition rates of students in the STEM fields continue to be a point of discussion. Combined with the demand for expansion in the STEM workforce, various retention reforms have been proposed, implemented, and in some cases integrated into policy following evidence of success. Still, new findings, technological advances, and socio-cultural shifts inevitably necessitate an ongoing investigation as to how students approach learning. Among other factors, students who enter college without effective study skills are at much greater risk of being unsuccessful in their coursework. In order to construct an equitable learning environment, a mechanism must be developed to provide underprepared students with access to resources or interventions designed to refine the skills they need to be successful in the course. Early, reliable assessments can provide predictions of individual student outcomes in order to guide the development and implementation of such targeted interventions. In the present study, a model is developed to predict students' odds of success based on their study approaches, as measured by their responses to twelve survey items from an existing instrument used in the Chemistry Education Research literature designed to measure students' deep and surface learning approaches. The model's prediction specificity ranges from 66.5% to 86.9% by semester. Two distinct sets of lower-performing students are identified in the data: those who align predominantly with surface approaches to learning versus those who indicate using both deep and surface approaches to learning. This supports the idea of a tailored approach to interventions, rather than a one-size-fits-all solution. Results from this instrument were correlated to students' reported study methods and beliefs.

KEYWORDS: *First-Year Undergraduate/General, Chemical Education Research, Testing/Assessment, Learning Theories*

FEATURE: Chemical Education Research

■ INTRODUCTION

Educators and researchers alike have sought to ameliorate the attrition rates and "weed out" connotation of the STEM gateway (or gatekeeping) courses. Potential solutions to these problems have included placement exams and/or remedial coursework; however, these measures may introduce financial burdens, time constraints, and other barriers that disproportionately impact students of nontraditional or marginalized status. Instead, the present study looks at the use of a the Modified Approaches and Study Skills Inventory (M-ASSIST),[1] to make predictions about students' course outcomes. The items from this instrument target students' study approaches, classifying them as deep or surface approaches. Combined with data collected on students' specific learning and study methods (e.g., attending lecture, reading the textbook), this research provides an imperfect but significant predictor of student outcomes. Such an instrument has the potential to provide instructors with the information needed to identify the distinct skills or approaches that at-risk students lack, rendering a more tailored approach to intervention possible.

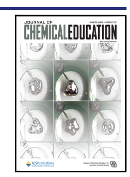
The following section will provide a background on some of the current practices in approaching the attrition problem, as well as previous efforts that have been taken to predict student outcomes and define deep and surface learning. A description of the setting for this study and the guiding research questions and methods will follow. Results are separated by research question, and a discussion section addresses the ways that these key findings are situated in the current literature. We conclude with a few general takeaways and implications for practitioners, along with an acknowledgment of limitations and future points of interest.

■ BACKGROUND

Placement, Interventions, and Equity

In the education literature, the term "placement" usually refers to the directing of students into a prerequisite[2−6] or corequisite course[7−9] that is deemed commensurate to their level of preparedness. Such placements have produced mixed results in the literature. While an online preparatory course at UC—Davis benefited underprepared students,[6] a multiyear study of another preparatory chemistry course at Texas Tech University concluded that the remediation provided "little or no significant academic benefit."[2] "Intervention", on the other

Received: August 12, 2020
Revised: November 17, 2020
Published: December 29, 2020



https://dx.doi.org/10.1021/acs.jchemed.0c01074
J. Chem. Educ. 2021, 98, 281−292

hand, typically refers to ancillary programs or activities within a course that aim to improve student outcomes with respect to specific course content (e.g., acids and bases[10]), skills (e.g., language comprehension[11]), or beliefs (e.g., growth mindset[12]). Benefits of early interventions in the classroom have been well-documented in first-year STEM courses.

The present study looks at students in the General Chemistry course sequence (GC1 and GC2) at Rutgers University, in which approximately one-quarter of the students earn grades of a D or F in the class (excluding students who withdraw). Students who do not perform well on the first exam in GC1 are strongly encouraged to switch into the Chemistry Preparatory (ChemPrep) course for the remainder of the semester. These students do not receive a "W" on their transcript for GC1, and they begin with a "clean slate" (gradewise) in the new course. Mills et al. describe a similar system after finding a high correlation with first exam performance and course grades.[13] While ChemPrep has anecdotal accounts of success, it is not without limitations. Not all students' schedules can accommodate a midsemester swap, which also places students at least one semester behind with few options for recovery. Summer coursework can prove impossible for students who do not live nearby, who lack the financial means, or who must spend this time working or tending to family. Alternatively, waiting until the fall postpones enrollment in subsequent courses such as Organic Chemistry, potentially delaying graduation and proving a financial burden.

While placements ensure that students do not become overwhelmed by material they are unprepared for, the very act of placement is inherently inequitable to students with financial insecurity or disabilities or who are part of marginalized communities. These students already face significant barriers when entering these academic spaces and leave at higher rates.[14] By identifying predictive factors of success, researchers and practitioners can work toward early, concurrent intervention where the goal is to retain students via a personalized approach, as opposed to placement. Effective interventions can then be sustained by incorporating course exam data as well as reassessing students' study habits across the academic term.

Predicting Success

Many studies have quantified students' odds for success and persistence in higher education. In the STEM education literature, factors linked to student outcomes include SAT scores,[15−17] GPA,[16,18,19] demographics,[16−18,20] and self-efficacy.[21] Content-based assessments such as the California Chemistry Diagnostic Exam[22] or the Toledo Chemistry Placement Exam[23] have used students' incoming content knowledge to predict outcomes. Not only have these efforts provided valuable information about a student's likelihood of success in courses, but they have also informed teaching practices and highlighted issues of equity in the classroom.

Another area of interest in terms of course outcomes is students' choice of study methods and the specific ways they employ these methods. In one investigation, Ye et al. used text messages to collect data on the types of study materials and frequency of use in a General Chemistry course.[24] In addition to linking study methods to outcomes, the authors found evidence that students changed their study methods over time, positing that recent exam content may have been the cause. In a second study by Ye et al., qualitative analysis suggested that the quality of studying was linked to at-risk students' course outcomes.[15] For example, several students reported studying with friends, but while some saw this as an opportunity to learn through teaching their peers ("deep approach"), another stated that they relied on their peers to help them or provide answers ("surface approach"). The current study uses some metrics to quantify the quality of studying and draws upon a similar deep/surface dichotomy, for the purpose of developing a predictive model of student success.

Deep and Surface Learning

The Approaches and Study Skills Inventory for Students (ASSIST) was developed by Tait et al. in 1997 and assesses students on their ideas about learning, study habits, and teaching preferences, classifying them as deep, strategic, or surface learners.[25,26] A shortened, modified version of this instrument, the M-ASSIST, was constructed by Bunce et al, in 2017, and examined deep and surface study approaches of General Chemistry students at the United States Naval Academy.[1] The authors define deep learners as those who purposefully attempt to connect new knowledge to that which they already know using the underlying concepts. In contrast, surface learners approach new knowledge in an algorithmic fashion, looking predominantly at the surface features of a problem and relying on rote memorization. The results showed that student success was positively correlated with deep study approaches and negatively correlated to surface study approaches.

In the present study, an investigation of such deep and surface learning approaches is used to construct a predictive model for student outcomes in General Chemistry. Identifying at-risk students at various points during the course may facilitate intervention over placement, while knowledge gained about students' learning approaches and habits may prove useful to instructors in determining the type of intervention needed for different students and at different times.

■ RESEARCH QUESTIONS

The first goal for the present study was to determine if the results from the M-ASSIST study could be replicated with a new population. Specifically, the M-ASSIST was examined as a potential predictive tool to identify at-risk students early on in the course. Further relating these deep and surface study approaches to specific habits (e.g., reading the textbook) may provide tangible advice or intervention strategies for these students. The research questions (RQs) pertinent to this study are as follows:

(1) To what extent can students' deep and/or surface approach(es) to studying, combined with demographic information, predict student success in general chemistry at a large, diverse, research-intensive institution?

(2) How do students' study habits correlate with their deep and surface study approaches as measured by the M-ASSIST?

■ SETTING

Population and Course Structure

The General Chemistry courses at Rutgers consist of large-enrollment lectures and weekly online, synchronous recitations that focus on problem-solving for topics covered during previous lectures. There are five sections of General Chemistry I each fall semester and four sections of General Chemistry II each spring. While each section is typically taught by a different instructor, the format of the lecture is the same for all sections, and students have access to any of the instructors' notes via the online course management system. Weekly homework is provided online via an in-house program with a combination of static and dynamic content, and students take three common-hour midterm exams, with both multiple-choice and open-ended components. The final exam consists of the most recent multiple-choice single-semester ACS exam plus five two-part open-ended questions. Teaching interns (TIs) hold optional supplemental instruction sessions, including workshops, office hours, and a review session.

■ DATA COLLECTION AND ANALYSIS

Modified Approaches and Study Skills Inventory (M-ASSIST)

In this study, the M-ASSIST was issued to students online during the first (pretest) and last (post-test) weeks of the Fall 2018, Spring 2019, and Fall 2019 semesters using Qualtrics. However, the remainder of this paper focuses on the post-test results of the M-ASSIST since the bulk of students' grades are determined in the final few weeks of the semester

The decision to administer the M-ASSIST was driven by practicality of implementation and its content agnosticism. The brevity and ease of scoring made it an attractive model to use in a class of 1500+ students. Further, the purpose was not to assess chemistry knowledge, and the researchers believe that the items on the M-ASSIST can be reasonably answered by students regardless of their chemistry background. It contains only twelve items of one sentence each, with six items contributing to the deep scale and six items to the surface scale, for which students are asked to note their level of agreement on a five-point Likert scale, specifically in the context of General Chemistry. Data was analyzed in R and SPSS (Version 26). Deep and surface scores are calculated by taking the average score of each subscale. Students who did not answer more than one item on both subscales were excluded from the analysis. The full M-ASSIST can be found in the original paper by Bunce et al.[1]

Student Individuality Survey (SIS)

A second survey, the Student Individuality Survey (SIS), was developed in-house and consists of two portions: The first asks students to provide demographic data, as well as course goals, and the extent of their previous high school and college chemistry coursework (Table 1). The second part of the survey includes a series of questions about students' learning and studying habits in the context of General Chemistry. The SIS was administered to students online alongside the M-ASSIST. A copy of this instrument can be found in the Supporting Information. SIS data was analyzed using SPSS Version 26.

Regression Analysis

Logistic regression was performed in the statistical program R using student outcomes as the dependent variable and various combinations of students' deep scores, surface scores, and demographics as the predictor variables. To determine the best model, the proposed models were compared using the Akaike's An Information Criterion (AIC), a multimodel inference technique.[27] In brief, the AIC value estimates the relative strength of each model within a set based on parsimony and goodness of fit. The smaller the AIC number within a set, the better the combination of fit and parsimony.

The ΔAIC is calculated by identifying the model with the smallest AIC and computing the absolute value of the difference between that model and all other models.[28] Only models in which all predictor variables (e.g., gender) have at least one significant individual factor component (e.g., female) are considered. The model with the smallest ΔAIC, which also meets this significance criterion, is selected. Further details on the statistical analyses and sample R commands are provided in the Supporting Information.

Table 1. Demographic Data for General Chemistry I (Fall 2018), N = 1455

Demographic Category	%
Gender	
Female	60.5%
Male	38.9%
Nonbinary/other	0.6%
Generation Status	
First-generation college student	27.4%
Major/Track[a]	
Life sciences	63.1%
Physical sciences	15.7%
Pharmacy	10.9%
Social science	6.2%
Engineering	3.0%
Other	1.1%
Prehealth track[b]	83.6%
Course Goals	
Earn an A	80.3%
Earn a B	18.0%
Earn a C/pass	1.7%
Race	
South/East Asian	45.8%
White	29.3%
Black/African-American	6.8%
Middle Eastern/North African	4.8%
Native Hawaiian/Pacific Islander	0.6%
Native American	0.3%
Two or more races	5.8%
Ethnicity	
Hispanic/Latino	12.4%
Previous Chemistry Coursework	
High school—none	1.1%
High school—1 semester to 1 year	66.7%
High school—2+ Years	31.2%
College—none	85.1%
College—1 semester	11.2%
College—2+ semesters	3.6%

[a]Major/career track data is based on the Fall 2015 cohort, as data for the Fall 2018 was unavailable [b]Students on the prehealth track may select any major(s)

IRB Approval and Consent Procedures

All methods and procedures were granted IRB approval from the institution, under IRB protocol 15-814M, with annual renewal.

■ RESULTS

RQ1, Part I: Defining Success

Students' study skill scores were measured on the deep and surface subscales and separated according to their final grade in the class. Figure 1 shows the distribution of average deep and surface scores for each letter group in the fall and spring semesters. Note that the surface score appears to be more

https://dx.doi.org/10.1021/acs.jchemed.0c01074
J. Chem. Educ. 2021, 98, 281−292

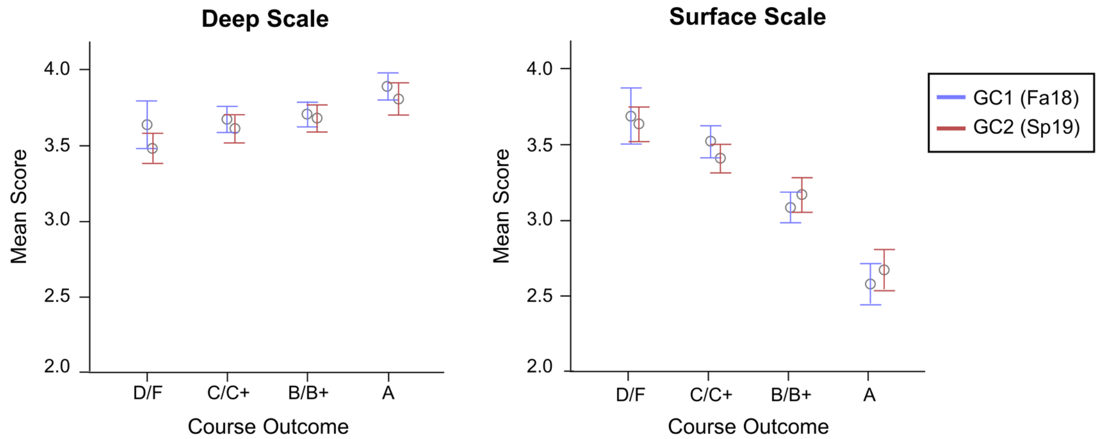


Figure 1. Average deep and surface scores from the M-ASSIST[1] were calculated according to the four grade groups for each semester and plotted on the graphs. Error bars represent the 95% confidence intervals.

sensitive than the deep score. This is consistent with the findings of by Bunce and colleagues.[1]

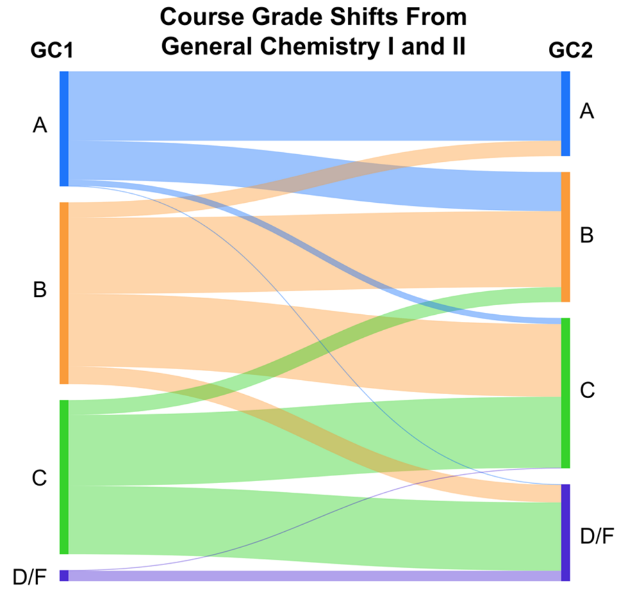
Gellene and Bentley suggest that multivariable prediction models perform best when the student outcome is binary.[2] In lieu of letter grades, student outcomes were labeled "successful" (S) or "unsuccessful" (U), with success defined as earning a grade of B or higher. The decision to use this cutoff stemmed from a few considerations. First, in both semesters, over 98% of the students selected a grade of "A" or "B" (Table 1) as their goal. Most convincingly, however, were the trends in grades from GC1 to GC2, illustrated in Figure 2. Of the students who earned an A in GC1, 94.3% of them earned a grade of A or B/B+ in GC2. Just over half of the students earning a B/B+ in GC1 earned a grade of A or B/B+ in GC2. Comparatively, not a single student from this cohort received an A in GC II following a grade of C/C+ in GC1, and only 9.6% of them earned a B/B+. This sharp contrast between the two groups lends support to the use of "B or better" as a demarcation line for success.

RQ1, Part II: Calculating Study Skills Scores

Figure 3 provides a breakdown of responses for each item on the M-ASSIST by outcome group for the Fall 2018 semester. An independent t test is used to investigate any differences between the two groups, and an effect size is calculated using Cohen's d. As a whole, there are greater differences between the two groups on the surface scale compared to the deep scale for both semesters, reflecting the findings from Figure 1.

For the fall term, analysis of the individual items yielded significant differences with moderate effect sizes for all items on the surface scale (Table 2). Items S2 and S6 both target sense-making and have the largest effect size. While the deep scale contains three items that suggest significant differences between the U/S groups, the data is underpowered to make definitive claims. The spring semester (Table S2, Supporting Information) followed a similar trend with respect to the surface scale; however, four items on the deep subscale were significantly different and achieved a statistical power of $\beta \geq 0.80$. Still, the effect sizes were considerably smaller compared to those of the surface scale.

Heatmaps were created by plotting each student according to their average deep and surface scores (Figure 4). Data points are colored on a gradient according to the proportion of successful students at that point. Areas with a higher proportion are coded in blue, while those with a lower proportion are coded in red. Whole-class average deep and surface scores form the four quadrants. In both semesters,



		Grade in GC1 (Fa18)			
		A	B	C	D/F
Grade in GC2 (Sp19)	A	60.3%	8.5%	0.0%	0.0%
	B	34.0%	41.8%	9.6%	0.0%
	C	5.3%	40.0%	46.1%	5.0%
	D/F	0.4%	9.7%	44.3%	95.0%
	SUM	100.0%	100.0%	100.0%	100.0%

Figure 2. (top) Sankey diagram showing the flow of grades students earned in GC1 (left) and then in GC2 (right). The width of the bands is proportional to the number of students represented. This data only shows students who completed GC1 in Fall 2018 and transitioned into and completed GC2 in Spring 2019 (N = 839). A full set of outcomes can be found in the Supporting Information (Table S1). For clarity, the percentages representing each band are provided in the grid (bottom) and are calculated as a percentage of students earning a given letter grade in GC1.

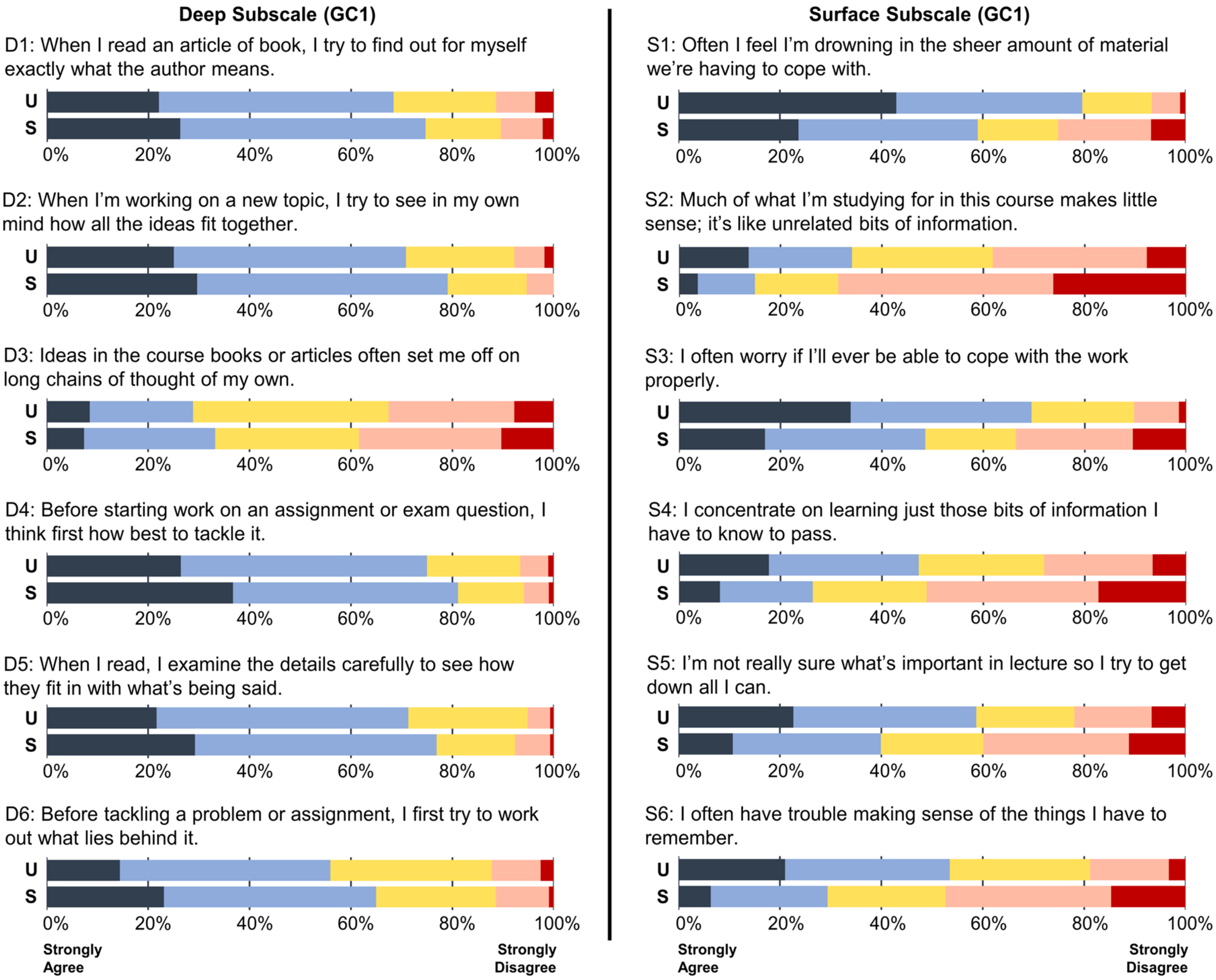


Figure 3. Responses from successful (S) and unsuccessful (U) students on the twelve items from the M-ASSIST for the Fall 2018 GC1 course. Responses of "strongly agree" are represented by dark blue (left side of scale), and responses of "strongly disagree" are in dark red (right side of scale). N_S = 411; N_U = 273. Responses for the Spring 2019 semester can be found in the Supporting Information.

Table 2. Differences per M-ASSIST Item for Successful and Unsuccessful Students[a] (Fall 2018)

	Deep (D)			Surface (S)		
Item	Sig.	Effect Size	Power	Sig.	Effect Size	Power
1	0.103	NS	NS	0.000	0.576	≥0.999
2	0.016	0.193	0.695	0.000	0.706	≥0.999
3	0.486	NS	NS	0.000	0.603	≥0.999
4	0.014	0.192	0.691	0.000	0.547	≥0.999
5	0.130	NS	NS	0.000	0.448	≥0.999
6	0.009	0.206	0.747	0.000	0.701	≥0.999

[a] N_S = 411; N_U = 273; "NS" is "not significant".

Quadrant 1 (top-right) contains the greatest density of blue (successful) data points. These are the students with high deep scores and low surface scores. Students in Quadrant 4 (bottom-right) did not have many successful outcomes, despite having above-average deep scores, illustrating the differences in sensitivity between the deep and surface scales seen previously in Figure 1. Table 3 lists the average fraction of successful students per quadrant to supplement the visual representations of the data. Note that, in both semesters, the quadrants with below-average surface scores had the largest fractions of successful students.

RQ1, Part III: Modeling and Predicting Success

Logistic regression was carried out on the binary outcome data (successful versus unsuccessful) as a function of various combinations of students' surface scores, deep scores, and demographics (first-generation status, gender, and race/ethnicity). Due to sample size, the categories of race and ethnicity were combined, as has been common practice in previous studies.[16,20,29] Table 4 provides an overview of the proposed models along with the AIC and ΔAIC values.

Table 5 provides the regression parameters associated with each model listed in Table 4. Each of the β_n values are log-odds parameters. The categorical variables produce parameters whose log-odds are relative to one of the component factors. Each factor is assigned a label of 1 through m, where m is the total number of factors within that categorical variable. Using the criteria described previously in this paper, the best models selected for each semester are as follows:

https://dx.doi.org/10.1021/acs.jchemed.0c01074
J. Chem. Educ. 2021, 98, 281−292

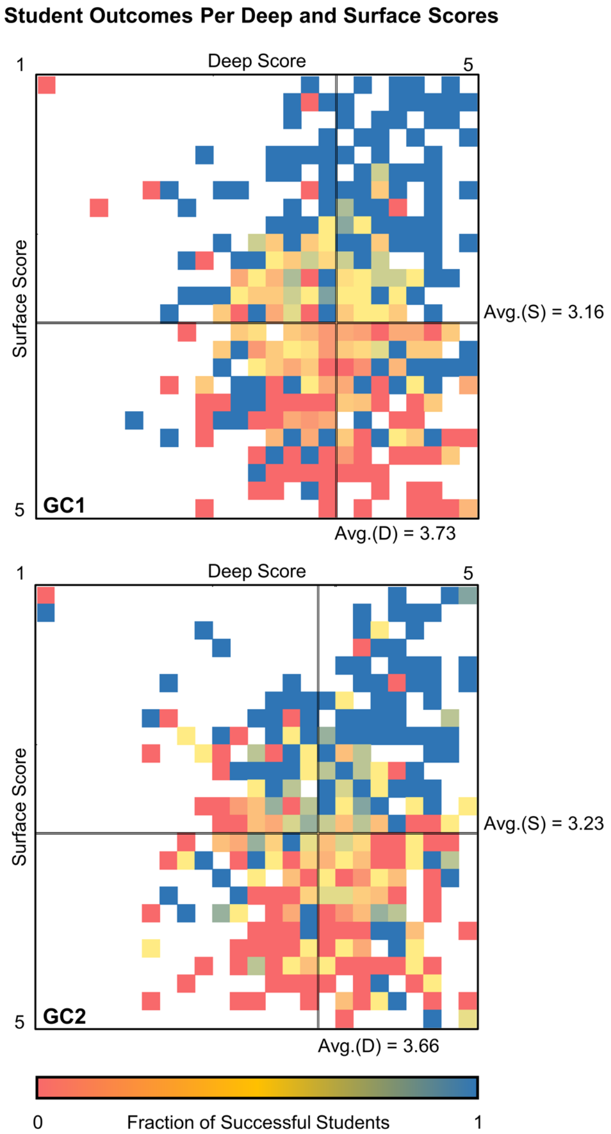


Figure 4. Heatmaps for the GC1 (top, $N = 653$) and GC2 (bottom, $N = 697$) courses were created by calculating the proportion of students that were successful at each possible combination of average deep (x-axis) and surface (y-axis) scores. Dark blue points represent a fraction of success = 1, whereas red represents a fraction of success = 0. White spaces indicate that no student had that combination of scores. Quadrants are formed using the overall average deep and surface scores and are numbered 1−4, starting in the top-right quadrant and proceeding counterclockwise.

Table 3. Fraction of Successful Students per Quadrant

			Fraction of Success	
Quadrant	Deep Scale	Surface Scale	GC1	GC2
1	High	Low	0.85	0.78
2	Low	Low	0.67	0.49
3	Low	High	0.43	0.38
4	High	High	0.41	0.35

$$\text{Outcome(Fall '18)} = \beta_0 + \beta_1\text{Deep} + \beta_2\text{Surface} + \beta_3\text{FirstGen} + \boldsymbol{\beta}_4\text{Race/Ethnicity} \quad (1)$$

$$\text{Outcome(Spring '19)} = \beta_0 + \beta_1\text{Deep} + \beta_2\text{Surface} + \boldsymbol{\beta}_4\text{Race/Ethnicity} \quad (2)$$

Table 4. Regression Models to Predict Student Outcomes in General Chemistry

Model	Predictor Variables	AIC	ΔAIC
	General Chemistry I—Fall 2018		
Fa0	deep + surface	811.0	76.4
Fa1	deep + surface + first generation	764.8	30.2
Fa2	deep + surface + first generation + gender	757.1	22.4
Fa3	deep + surface + first generation + gender + race/ethnicity	734.7	0.0
Fa4	deep + surface + gender + race/ethnicity	751.5	16.8
Fa5	deep + surface + race/ethnicity	760.4	25.7
Fa6	deep + surface + first generation + race/ethnicity	743.5	8.9
	General Chemistry II—Spring 2019		
Sp0	deep + surface	908.7	115.9
Sp1	deep + surface + first generation	821.1	28.3
Sp2	deep + surface + first generation + gender	816.6	23.8
Sp3	deep + surface + first generation + gender + race/ethnicity	792.8	0.0
Sp4	deep + surface + gender + race/ethnicity	804.0	11.2
Sp5	deep + surface + race/ethnicity	808.4	15.6
Sp6	deep + surface + first generation + race/ethnicity	796.8	4.0

The model labeled eq 1 is derived from the Fall 2018 data, while eq 2 is derived from the Spring 2019 data. Only the predictor variables with significant regression parameters are included in the best models. The bold $\boldsymbol{\beta}_4$ terms refer to a set of parameters related to the Race/Ethnicity variable. In the case of Model Fa6 (eq 1), the significant factor component for the Race/Ethnicity categorical variable is that for Asian (RaceEth3, Table 5). For Model Sp5 (eq 2), the significant factor component within the same variable is that for Hispanic/Latinx (RaceEth2, Table 5).

To evaluate the predictive capabilities of these two models, outcome probabilities are computed by plugging in students' data into the selected models. These probabilities are translated into predicted outcomes using the following decision boundary: a probability of ≥0.5 was assigned "1" (successful) while a probability of <0.5 was assigned "0" (unsuccessful). Students' predicted outcomes were then compared to their actual outcomes in the course (Table 6).

In the Fall 2018 semester, Model Fa6 correctly predicts an unsuccessful outcome in the course slightly less than 50% of the time but predicts successful outcomes slightly more than 80% of the time. This remains true even when the same model is tested with a different cohort in Fall 2019. GC2 Model Sp5 correctly predicts desirable and undesirable outcomes nearly two-thirds of the time. Results from testing the two models on the alternate semesters (i.e., Sp5 model used on GC1 data, Fa6 model used on GC2 data) are also provided in Table 6. In both cases, the use of these alternate models provides lower overall prediction rates, supporting the use of two different models, Fa6 and Sp5, for their respective semesters.

RQ2: Study Skills and Academic Habits

Students' lecture engagement habits were correlated with their average deep and surface scores using a Spearman rank order correlation (Table 7). In both GC1 and GC2, average deep scores are significantly correlated with all five items listed under lecture engagement habits, though the correlation coefficients were small in magnitude. The item "focus in lecture" has the largest correlation with the deep score in both semesters and the only significant correlation with the surface score, which is negative.

https://dx.doi.org/10.1021/acs.jchemed.0c01074
J. Chem. Educ. 2021, 98, 281−292

Table 5. Regression Model Parameters to Predict Student Outcomes in General Chemistry[a]

Model	Log-Odds Parameters (β_n) Intercept	Deep	Surface	FirstGen2	Gender2	RaceEth2	RaceEth3	RaceEth4
General Chemistry I—Fall 2018								
Fa0	2.364***	0.388**	−1.042***					
Fa1	2.582***	0.414**	−1.071***	−0.649**				
Fa2	2.345***	0.441**	−1.067***	−0.658**	0.324			
Fa3	2.124**	0.429**	−1.073***	−0.638**	0.358	−0.250	0.498*	0.402
Fa4	1.848**	0.430**	−1.051***		0.346	−0.366	0.542**	0.440
Fa5	2.112***	0.404**	−1.054***			−0.388	0.500*	0.390
Fa6	2.403***	0.400**	−1.076***	−0.637**		−0.271	0.459*	0.355
General Chemistry II—Spring 2019								
Sp0	1.136*	0.516***	−0.920***					
Sp1	1.699**	0.429**	−0.943***	−0.459*				
Sp2	1.470*	0.452***	−0.931***	−0.441*	0.265			
Sp3	1.612*	0.454**	−0.971***	−0.392	0.271	−0.787	0.090	−0.190
Sp4	1.366*	0.499***	−0.973***		0.263	−0.926*	0.086	−0.202
Sp5	1.602**	0.479***	−0.987***			−0.980*	0.065	−0.232
Sp6	1.851**	0.434**	−0.984***	−0.402		−0.829*	0.070	−0.220

[a]*$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$. FirstGen2, holds first-generation college status; Gender2, Male; RaceEth2, Hispanic/Latinx; RaceEth3, Asian; RaceEth4, Black or African-American

Table 6. Predictive Capabilities of Regression Models

Value	GC1—Fall 2018		GC2—Spring 2019		GC1—Fall 2019
Model used	Fa6[a]	Sp5	Sp5[a]	Fa6	Fa6
N_{Total}	643	649	652	644	485
N_U (actual)	248	252	312	307	203
N_S (actual)	395	397	340	337	282
Specificity[b]	80.0%	70.3%	66.5%	81.0%	86.9%
Sensitivity[c]	48.4%	61.1%	62.5%	44.6%	48.8%
% Pos. predictive value[d]	71.2%	74.0%	65.9%	61.6%	70.2%
% Neg. predictive value[e]	60.3%	56.6%	63.1%	68.2%	72.8%
% Predicted overall	67.8%	66.7%	64.6%	63.7%	70.9%

[a]Model selected based on ΔAIC. [b]% of successful outcomes correctly predicted by model. [c]% of unsuccessful outcomes correctly predicted by model. [d]% of successful predictions that were correct. [e]% of unsuccessful predictions that were correct.

Table 7. Spearman Correlations of Study Skills and Lecture Engagement Habits[a]

"How frequently do you do the following?" (5 Point Likert Scale)	GC1—Fall 2018 Deep	GC1—Fall 2018 Surface	GC2—Spring 2019 Deep	GC2—Spring 2019 Surface
Attend lecture	0.094*	−0.035	0.155***	−0.063
Prepare before lecture	0.178***	0.018	0.149***	0.016
Take notes during lecture	0.122***	0.014	0.104**	−0.027
Pay attention in lecture	0.208***	−0.154***	0.208***	−0.147***
Take notes while reading textbook/lecture notes	0.191***	0.048	0.169***	0.009

[a]*$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$.

Spearman correlations were also calculated for students' deep and surface study skills with their general approaches and beliefs toward studying (Table 8). Overall, this section encompasses the largest correlation coefficients. Satisfaction with study habits is positively correlated with the deep score in both semesters and negatively correlated with the surface score, suggesting that students do have some awareness of their academic progress in the class. However, results on the second item suggest that students with higher surface scores may not know how to improve their study habits.

On the SIS, cramming is defined as "mass studying in the last day or two before an exam, rather than spread out." The frequency of cramming for exams is positively correlated with the surface score and negatively correlated with the deep score for both semesters. Still, when those who indicate at least some tendency to cram are asked about the effectiveness of their cramming, no clear trend could be identified except for a small negative correlation with the spring's surface score.

The final item in this section asks students whether or not they felt they had to memorize a significant amount of material in the class. Agreement with this item resulted in a low negative correlation with the deep score and a moderate positive correlation with the surface score. Despite the different content material presented in GC1 and GC2, these results were consistent in both semesters and align with that which would be expected from the M-ASSIST.

Finally, students were provided with a set of eight study methods and told to select as many methods as they actually found helpful during the semester. Neither are the preferred study methods consistent between semesters, nor were any of the correlation coefficients sufficiently large (Table 9). The only study methods that produced significant results in both GC1 and GC2 were reading the textbook (positively correlated with the deep score) and watching videos online (positively correlated with the surface score). Online videos are not a component of the course and thus refer to any videos from third parties that students sought independently.

One possible explanation for the low predictability of at-risk (predicted-unsuccessful) students in the fall is that their study methods might vary. These students were divided into two groups: at-risk, successful (N = 76); and at-risk, unsuccessful (N = 117). The percentage of students in these two groups

https://dx.doi.org/10.1021/acs.jchemed.0c01074
J. Chem. Educ. 2021, 98, 281−292

Table 8. Spearman Correlations of Study Skills and Beliefs about Habits[a]

Frequency/Agreement with the Following	GC1—Fall 2018		GC2—Spring 2019	
	Deep	Surface	Deep	Surface
I am satisfied with my study habits (3 pt-Likert scale)	0.180***	−0.327***	0.112**	−0.335***
I know how to improve my study habits[b] (T/F)	0.078	−0.315***	0.112**	−0.159***
How often do you cram before exams? (5-pt Likert scale)	−0.176***	0.400***	−0.137***	0.246***
Do you believe cramming works well for you?[c] (3 pt Likert scale)	−0.003	−0.073	−0.015	−0.143***
I find myself having to memorize a significant amount of material in this class (T/F)	−0.243***	0.454***	−0.179***	0.391***

[a]*$p \le 0.05$; **$p \le 0.01$; ***$p \le 0.001$. [b]This item was only available for those who selected "Somewhat Satisfied" or "Not Satisfied" with the previous item, "I am satisfied with my study habits". [c]This item was only available for those who did not select "Never" to the previous item, "How often do you cram before exams?".

Table 9. Spearman Correlations of Study Skills and Study Habits[a]

"How helpful do you find the following when studying?" (3 Point Likert Scale)	GC1—Fall 2018		GC2—Spring 2019	
	Deep	Surface	Deep	Surface
Reading the textbook	0.080*	−0.027	0.151***	0.020
Reading the instructor's notes	0.007	−0.146***	0.042	−0.072
Reading another instructor's notes	−0.049	0.081*	−0.040	0.072
Watching videos online	0.033	0.219***	−0.027	0.132***
Writing own notes	0.034	−0.037	0.067	−0.039
Doing practice problems from the textbook	0.149***	−0.118**	0.056	0.001
Doing practice problems from outside of the textbook	−0.005	−0.013	0.024	−0.038
Redoing the homework	0.092*	−0.053	0.042	0.003

[a]*$p \le 0.05$; **$p \le 0.01$; ***$p \le 0.001$.

who utilize each of the study methods was calculated and compared using a Chi-square test (Table 10). For GC1 in Fall 2018, watching videos online is the only study method that shows a significant difference between the successful and unsuccessful students that were initially deemed at-risk.

Table 10. Study Methods of At-Risk Students by Course Outcomes (Fall 2018)[a]

Study Method	At-Risk (S) % Use	At-Risk (U) % Use	Chi-Square, X^2	Sig.
Reading the textbook	52.5	59.5	0.983	
Reading the instructor's notes	78.8	77.0	0.088	
Reading another instructor's notes	62.5	57.1	0.582	
Watching videos online	57.5	84.1	17.9	$p \le 0.001$
Writing own notes	65.0	61.1	0.316	
Doing practice problems from the textbook	56.3	56.4	0.000	
Doing practice problems from outside of the textbook	60.0	60.3	0.002	
Redoing the homework	32.5	42.1	1.89	

[a]$N_S = 76$; $N_U = 117$.

■ DISCUSSION

Deep and Surface Subscales

While students' surface scores on the M-ASSIST exhibit clear differences between the achievement groups, the deep scale appears to be less sensitive overall. These findings mimic those reported in the original M-ASSIST study by Bunce and colleagues,[1] which found that the surface scale could readily differentiate between the three grade groups (A/B, C, and D/F), while the deep scale did so to a lesser extent.

The relationships between students' deep and surface scores with their course outcomes are readily visualized by the heatmaps in Figure 4. Notably, the quadrants with below-average surface scores contain the largest fraction of students earning a B or better in the course. Specifically, students with below-average surface scores and above-average deep scores seem to fare the best. Interestingly, however, the same trend is not apparent with students who have higher surface scores. That is, the differences in the fractions of success between Quadrants 3 and 4 are less apparent than the differences between Quadrants 1 and 2. In fact, students in Quadrant 4 (high surface/high deep) as a whole had lower proportions of success than students than students in Quadrant 3 (high surface/low deep), for both semesters. While the deep scale was found to be less sensitive than the surface scale, that students in Quadrant 4 had lower proportions of success was unexpected.

Previous work on study approaches suggests that students within this quadrant are not homogeneous in terms of their beliefs and approaches to studying for coursework. Entwistle et al. used a cluster analysis to characterize student responses to the Approaches to Studying Inventory (ASI), the predecessor to the M-ASSIST.[30] While most clusters appeared to be typical (i.e., deep and surface scores were inversely related), one cluster reported unusually high deep scores with high surface scores. This particular cluster was the second-lowest in academic performance (out of six) and was not far behind the lowest-performing cluster.

Entwistle describes these high-deep/high-surface students as "disorganised in their studying, highly anxious and with confusion in ... their intention to seek meaning and declared interest in the ideas in the course, on the one hand, and their ... weak levels of understanding on the other."[30] He suggests a differentiation among the lower-performing students, specifically between the students with genuine surface approaches and those who are likely deep learners but who do not know how to properly utilize those approaches (and fall back on surface approaches instead). This dissonance is recognized in similar studies[31−33] and is consistent with what is known about metacognitive skills and course performance. Students who effectively utilize metacognitive strategies, such as evaluating their understanding and monitoring their study habits, tend to perform better academically.[15,34−37] Some surface learners may perform poorly simply due to their surface approaches to learning (failing to evaluate their understanding). Others may acknowledge that a deep approach is more effective but are unsure as to how to execute that approach successfully (failing to monitor their study habits).

Modeling Success

In each of the regression models for both the fall and spring semesters, the deep and surface scores emerged as the strongest predictors among the independent variables (Table 5). When incorporating the demographic data, identifying as Hispanic (GC2) and/or first-generation (GC1) was found to be negatively associated with student outcomes, while identifying as Asian (GC1) was found to be positively associated with student outcomes. These findings are consistent with the current literature[14,29,38] but still serve as an important reminder of the role of student identity in this research.

Overall, both the Fall 2018 and Spring 2019 models were able to correctly predict student outcomes roughly two-thirds of the time (Table 6). Gellene and Bentley estimate that, even with a binary outcome, the predictive accuracy of multivariable models reaches a maximum around 70−80% due to "intangible" quantities such as individual motivation.[2] The Fall 2018 GC1 model exhibited a large disparity in its specificity versus sensitivity (80.0% versus 48.4%, respectively), which was repeated when the model was applied to data from the Fall 2019 cohort. This sensitivity/specificity gap has been previously observed by other researchers, though a definitive explanation has not been established.[13,39] Still, the consistency between the two fall semesters was encouraging. The model applied to the GC2 course was more equitable in its predictions, correctly identifying the outcomes of successful students 66.5% of the time, and unsuccessful students 62.5% of the time. It is possible that the students in GC2 are a more homogeneous cohort due to a "filter effect" resulting from GC1-to-GC2 attrition. The degree to which such homogeneity accounts for the fidelity of the predictive models warrants further investigation.

Study Methods and Metacognition

The second research question focused on the habits and study methods that students report using in the class. Generally, favorable lecture engagement habits (e.g., preparing ahead of time) were positively correlated with the deep score (Table 7), while crammed studying and memorization of content were positively correlated with the surface score (Table 8).

Examination of the specific study methods that students report using in class indicates that deep and surface learners draw upon many of the same resources (Table 9). However, watching videos emerged as a practice moderately correlated with the surface score in both GC1 and GC2. Students who were deemed at risk and were ultimately unsuccessful were significantly more likely to report watching these videos compared to the at-risk, successful group (Table 10). Online media continues to mold the educational landscape and has offered many benefits to learning.[40] However, students who lack effective metacognitive skills may be more prone to passive learning at best and miscalibrated confidence at worst.[41−43] Students with high surface scores, who are thus predicted to be at-risk, may be more prone to using these unhelpful practices while engaging with online videos. This is particularly true if the online videos do not have built-in features to encourage students to reflect or self-assess on content related to the subject matter.

■ CONCLUSIONS

Drawing on one full year of General Chemistry at a large R1 university, a logistic regression model containing predictor variables of deep scores, surface scores, and demographic data has an overall prediction accuracy between 65% and 70%. Notably, the surface scores are the strongest predictors of success. It is a promising finding that the deep and surface scales' sensitivities were consistent with those found by Bunce and colleagues,[1] despite the fact that the present study's cohorts were quite different.

Although numerous placement tests have been previously described in detail, the M-ASSIST does have some unique benefits. First, the M-ASSIST consists of only 12 items, can easily be administered online, and typically takes less than 10 min to complete. Second, the M-ASSIST does not require any previous chemistry, math, or other STEM content knowledge. Lastly, the M-ASSIST can be quickly scored by instructors using any type of data analysis software or spreadsheet program and can provide actionable feedback for students if serving as an advisory tool.

It was unfortunately timely that this study coincided in part with the peak of the SARS-CoV2 pandemic. Although not the intention, these circumstances serve as a reminder of how important it is for students to develop effective, independent study methods and approaches. General Chemistry is typically taken by students in their first year of college, while they are adjusting to a new setting, new responsibilities, and new freedoms. Compounded with poor metacognitive skills and an unlimited amount of resources at their disposal, some students may see these introductory courses as an obstacle to overcome, rather than as a stepping-stone toward their goals.

■ IMPLICATIONS FOR INSTRUCTION

The results from this study have precipitated three main implications. The first is that deep and surface learning approaches, as measured by the M-ASSIST, do not necessarily exist on a single spectrum, and thus, students' placement on the surface scale, for example, may not be related to their placement on the deep scale. This suggests that a one-size-fits-all solution may not be suitable. For example, while studies have reported positive outcomes following in-class interventions on metacognitive strategies,[44,45] one study found that high-achieving students may actually have adverse reactions to this type of intervention.[46] Instead, a prediction model that incorporates results from the M-ASSIST administered across multiple parts of the academic term may be useful to test and ultimately identify the best intervention for different students. It is worth noting that examining how students' study skills evolve over the course of a term could help fine-tune such interventions (i.e., an intervention for a student at the beginning of the term may not be appropriate for that same student as their study skills and attitudes have possibly evolved later in the term).

https://dx.doi.org/10.1021/acs.jchemed.0c01074
J. Chem. Educ. 2021, 98, 281−292

Second, results suggest that successful and unsuccessful students in this cohort do not appear to use drastically different study methods from one another. This serves as a reminder of the language gap between students and instructors, which can be succinctly summarized by Cook and collaborators:[44]

> *... when students learn about Bloom's taxonomy, which almost none of them have seen before, they understand what faculty members mean by higher-order thinking. If students have never been explicitly taught that there is more to learning than memorization, they have no way of knowing how to develop higher-order thinking skills.*

The authors here argue that vague phrases like "higher-order thinking" are not helpful for students who do not know how to apply these ideas in a tangible way. Analogously, one-dimensional or cliched study advice like "don't cram" or "read the textbook" not only make assumptions about an individual's prior knowledge about learning but also ignore factors that might place them at risk. Instead, students in need of studying assistance should be guided in developing specific, actionable measures that they can reasonably implement. Instructors should avoid vague advice and be cognizant of the different ways that students utilize a given study method, as some may result in unproductive or deleterious outcomes.

Finally, one notable finding was that, for at-risk, unsuccessful students, a higher frequency of studying via online videos is reported. There is no dearth of best practices literature on the use of videos education, such as the use of guiding questions or interspersed polling.[47−50] However, in the case of third-party videos that students seek independently, instructors and peer leaders should take time to educate (and remind) students of how to properly use these videos and monitor their understanding, emphasizing the pitfalls of passive learning or false confidence.[51]

■ FUTURE DIRECTIONS AND LIMITATIONS

This paper describes the first use of the M-ASSIST as a means for predicting student success in General Chemistry. The majority of students in this cohort were life-science majors with an interest in health professional careers. Further work could investigate how these study skills may differ among engineers, chemistry majors, and a variety of other student cohorts at different types of institutions (e.g., small liberal arts colleges, minority serving institutions, and regional comprehensive colleges and universities). The research team is also interested in eventually expanding this research to include students at the organic chemistry level as well as students in off-sequence courses. Moreover, additional work could be done to provide evidence toward the validity and reliability of the M-ASSIST in various populations.

Further, there were no clear indications as to why the surface scale was more sensitive in predicting student outcomes although differences in students' metacognitive skills may be implicated. Likewise, students across the spectra of study skills and outcomes generally reported using similar study methods on their own. The next step may be to investigate the ways that students actually engage with these resources. Modifications to the SIS might help to move past the "what" into the "how" and would be better informed by qualitative data, such as interviews or focus groups. Such work could also increase the degree to which we can move students away from relying on surface-level strategies to approach their coursework and toward more deep, meaningful, and research informed methods.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available at https://pubs.acs.org/doi/10.1021/acs.jchemed.0c01074.

Additional statistical information with sample R commands, supporting data that was collected, and a copy of the Student Individuality Survey (SIS) instrument (PDF, DOCX)

■ AUTHOR INFORMATION

Corresponding Author

Marc N. Muñiz − *Department of Chemistry and Chemical Biology and Cyberlearning Innovation and Research Center, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854-8087, United States;* orcid.org/0000-0002-0270-3012; Email: mnm111@chem.rutgers.edu

Authors

Emily L. Atieh − *Department of Chemistry and Chemical Biology and Cyberlearning Innovation and Research Center, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854-8087, United States;* orcid.org/0000-0002-7143-1209

Darrin M. York − *Department of Chemistry and Chemical Biology and Cyberlearning Innovation and Research Center, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854-8087, United States;* orcid.org/0000-0002-9193-7055

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jchemed.0c01074

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We would like to express our sincerest gratitude to the Rutgers Department of Chemistry and Chemical Biology for their support in this project, including the General Chemistry instructors and students for their cooperation in conducting this research. We would also like to thank Mary Emenike for her insight on this project. This work would not have been possible without the support, infrastructure, and collaboration of the Cyberlearning Innovation and Research Center, funded by the School of Arts and Sciences at Rutgers, New Brunswick. Lastly, we would like to acknowledge funding provided by the National Institutes of Health [GM62248].

■ REFERENCES

(1) Bunce, D. M.; Komperda, R.; Schroeder, M. J.; Dillner, D. K.; Lin, S.; Teichert, M. A.; Hartman, J. R. Differential Use of Study Approaches by Students of Different Achievement Levels. *J. Chem. Educ.* **2017**, *94* (10), 1415−1424.

(2) Gellene, G. I.; Bentley, A. B. A Six-Year Study of the Effects of a Remedial Course in the Chemistry Curriculum. *J. Chem. Educ.* **2005**, *82* (1), 125−130.

(3) Heredia, K.; Xu, X. Y.; Lewis, J. E. The Application and Evaluation of a Two-Concept Diagnostic Instrument with Students Entering College General Chemistry. *Chem. Educ. Res. Pract.* **2012**, *13* (1), 30−38.

(4) Donovan, W. J.; Wheland, E. R. Comparisons of Success and Retention in a General Chemistry Course before and after the

Adoption of a Mathematics Prerequisite. *Sch. Sci. Math.* **2009**, *109* (7), 371–382.
(5) Freeman, W. A. Relative Long-Term Benefits of a Psi and a Traditional-Style Remedial Chemistry Course. *J. Chem. Educ.* **1984**, *61* (7), 617.
(6) Dockter, D.; Uvarov, C.; Guzman-Alvarez, A.; Molinaro, M. Improving Preparation and Persistence in Undergraduate Stem: Why an Online Summer Preparatory Chemistry Course Makes Sense. In *Online Approaches to Chemical Education*; ACS Publications, 2017; pp 7–33.
(7) Pickering, M. Helping the High Risk Freshman Chemist. *J. Chem. Educ.* **1975**, *52* (8), 512–514.
(8) Shields, S. P.; Hogrebe, M. C.; Spees, W. M.; Handlin, L. B.; Noelken, G. P.; Riley, J. M.; Frey, R. F. A Transition Program for Underprepared Students in General Chemistry: Diagnosis, Implementation, and Evaluation. *J. Chem. Educ.* **2012**, *89* (8), 995–1000.
(9) Shah, L.; Butler Basner, E.; Ferraro, K.; Sajan, A.; Fatima, A.; Rushton, G. T. Diversifying Undergraduate Chemistry Course Pathways to Improve Outcomes for at-Risk Students. *J. Chem. Educ.* **2020**, *97* (7), 1822–1831.
(10) Cox, C. T.; Poehlmann, J. S.; Ortega, C.; Lopez, J. C. Using Writing Assignments as an Intervention to Strengthen Acid-Base Skills. *J. Chem. Educ.* **2018**, *95* (8), 1276–1283.
(11) Pyburn, D. T.; Pazicni, S.; Benassi, V. A.; Tappin, E. M. The Testing Effect: An Intervention on Behalf of Low-Skilled Comprehenders in General Chemistry. *J. Chem. Educ.* **2014**, *91* (12), 2045–2057.
(12) Fink, A.; Cahill, M. J.; McDaniel, M. A.; Hoffman, A.; Frey, R. F. Improving General Chemistry Performance through a Growth Mindset Intervention: Selective Effects on Underrepresented Minorities. *Chem. Educ. Res. Pract.* **2018**, *19* (3), 783–806.
(13) Mills, P.; Sweeney, W.; Bonner, S. M. Using the First Exam for Student Placement in Beginning Chemistry Courses. *J. Chem. Educ.* **2009**, *86* (6), 738–743.
(14) Harris, R. B.; Mack, M. R.; Bryant, J.; Theobald, E. J.; Freeman, S. Reducing Achievement Gaps in Undergraduate General Chemistry Could Lift Underrepresented Students into a "Hyperpersistent Zone". *Science Advances* **2020**, *6* (24), No. eaaz5687.
(15) Ye, L.; Shuniak, C.; Oueini, R.; Robert, J.; Lewis, S. Can They Succeed? Exploring at-Risk Students' Study Habits in College General Chemistry. *Chem. Educ. Res. Pract.* **2016**, *17* (4), 878–892.
(16) Crisp, G.; Nora, A.; Taggart, A. Student Characteristics, Pre-College, College, and Environmental Factors as Predictors of Majoring in and Earning a Stem Degree: An Analysis of Students Attending a Hispanic Serving Institution. *Am. Educ. Res. J.* **2009**, *46* (4), 924–942.
(17) Tai, R. H.; Sadler, P. M.; Loehr, J. F. Factors Influencing Success in Introductory College Chemistry. *J. Res. Sci. Teach.* **2005**, *42* (9), 987–1012.
(18) Eddy, S. L.; Brownell, S. E.; Wenderoth, M. P. Gender Gaps in Achievement and Participation in Multiple Introductory Biology Classrooms. *CBE-Life Sci. Educ.* **2014**, *13* (3), 478–492.
(19) Griffith, A. L. Persistence of Women and Minorities in Stem Field Majors: Is It the School That Matters? *Econ. Educ. Rev.* **2010**, *29* (6), 911–922.
(20) Shaw, E.; Barbuti, S. Patterns of Persistence in Intended College Major with a Focus on Stem Majors. *NACADA J.* **2010**, *30*, 19.
(21) Pintrich, P. R.; Smith, D. A. F.; Garcia, T.; McKeachie, W. J. Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire (Mslq). *Educ. Psychol. Meas.* **1993**, *53* (3), 801–813.
(22) Legg, J. C.; Greenbowe, T. J.; Legg, M. J. Analysis of Success in General Chemistry Based on Diagnostic Testing Using Logistic Regression. *J. Chem. Educ.* **2001**, *78* (8), 1117.
(23) Hovey, N. W.; Krohn, A. An Evaluation of the Toledo Chemistry Placement Examination. *J. Chem. Educ.* **1963**, *40* (7), 370–372.
(24) Ye, L.; Oueini, R.; Dickerson, A. P.; Lewis, S. E. Learning Beyond the Classroom: Using Text Messages to Measure General Chemistry Students' Study Habits. *Chem. Educ. Res. Pract.* **2015**, *16* (4), 869–878.
(25) Tait, H.; Entwistle, N.; McCune, V. Assist: A Reconceptualisation of the Approaches to Studying Inventory. In *Improving Students as Learners*; Rust, C., Ed.; Oxford Brooks University, Oxford Centre for Staff and Learning Development: Oxford, 1998; pp 262–271.
(26) Entwistle, N.; McCune, V.; Tait, H. *Approaches and Study Skills Inventory for Students (Assist): Report of the Development and Use of the Inventories*; Unpublished.
(27) Akaike, H. A New Look at the Statistical Model Identification. *IEEE Trans. Autom. Control* **1974**, *19* (6), 716–723.
(28) Burnham, K. P.; Anderson, D. R. Multimodel Inference: Understanding Aic and Bic in Model Selection. *Sociol. Method. Res.* **2004**, *33* (2), 261–304.
(29) Fink, A.; Frey, R. F.; Solomon, E. D. Belonging in General Chemistry Predicts First-Year Undergraduates' Performance and Attrition. *Chem. Educ. Res. Pract.* **2020**, *21*, 1042.
(30) Entwistle, N.; Tait, H.; McCune, V. Patterns of Response to an Approaches to Studying Inventory across Contrasting Groups and Contexts. *Eur. J. Psychol. Educ.* **2000**, *15* (1), 33–48.
(31) Meyer, J. H. F. The Modelling of 'Dissonant' Study Orchestration in Higher Education. *Eur. J. Psychol. Educ.* **2000**, *15* (1), 5–18.
(32) Meyer, J. H. F. Study Orchestration: The Manifestation, Interpretation and Consequences of Contextualised Approaches to Studying. *High. Educ.* **1991**, *22* (3), 297–316.
(33) Lindblom-Ylänne, S.; Lonka, K. Dissonant Study Orchestrations of High-Achieving University Students. *Eur. J. Psychol. Educ.* **2000**, *15* (1), 19–32.
(34) Rickey, D.; Stacy, A. M. The Role of Metacognition in Learning Chemistry. *J. Chem. Educ.* **2000**, *77* (7), 915.
(35) Kruger, J.; Dunning, D. Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments. *J. Pers. Soc. Psychol.* **1999**, *77* (6), 1121–1134.
(36) Pintrich, P. R. The Role of Metacognitive Knowledge in Learning, Teaching, and Assessing. *Theor. Pract.* **2002**, *41* (4), 219–225.
(37) Tanner, K. D. Promoting Student Metacognition. *CBE-Life Sci. Educ.* **2012**, *11* (2), 113–120.
(38) MacPhee, D.; Farro, S.; Canetto, S. S. Academic Self-Efficacy and Performance of Underrepresented Stem Majors: Gender, Ethnic, and Social Class Patterns. *Anal. Soc. Iss. Pub. Pol.* **2013**, *13* (1), 347–369.
(39) Wagner, E. P.; Sasser, H.; DiBiase, W. J. Predicting Students at Risk in General Chemistry Using Pre-Semester Assessments and Demographic Information. *J. Chem. Educ.* **2002**, *79* (6), 749.
(40) Means, B.; Toyama, Y.; Murphy, R.; Bakia, M.; Jones, K. *Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies*; U.S. Department of Education: Washington, D.C., 2009.
(41) Szpunar, K. K.; Jing, H. G.; Schacter, D. L. Overcoming Overconfidence in Learning from Video-Recorded Lectures: Implications of Interpolated Testing for Online Education. *J. Appl. Res. Mem. Cogn.* **2014**, *3* (3), 161–164.
(42) Schacter, D. L.; Szpunar, K. K. Enhancing Attention and Memory During Video-Recorded Lectures. *Scholarsh. Teach. Learn. Psychol.* **2015**, *1* (1), 60–71.
(43) Finn, B.; Tauber, S. K. When Confidence Is Not a Signal of Knowing: How Students' Experiences and Beliefs About Processing Fluency Can Lead to Miscalibrated Confidence. *Educ. Psychol. Rev.* **2015**, *27* (4), 567–586.
(44) Cook, E.; Kennedy, E.; McGuire, S. Y. Effect of Teaching Metacognitive Learning Strategies on Performance in General Chemistry Courses. *J. Chem. Educ.* **2013**, *90* (8), 961–967.
(45) Mutambuki, J. M.; Mwavita, M.; Muteti, C. Z.; Jacob, B. I.; Mohanty, S. Metacognition and Active Learning Combination Reveals Better Performance on Cognitively Demanding General Chemistry

https://dx.doi.org/10.1021/acs.jchemed.0c01074
J. Chem. Educ. 2021, 98, 281–292

Concepts Than Active Learning Alone. *J. Chem. Educ.* **2020**, *97* (7), 1832−1840.

(46) Ee, J.; Moore, P. J.; Atputhasamy, L. High-Achieving Students: Their Motivational Goals, Self-Regulation and Achievement and Relationships to Their Teachers' Goals and Strategy-Based Instruction. *High Abil. Stud.* **2003**, *14* (1), 23−39.

(47) Petillion, R. J.; McNeil, W. S. Johnstone's Triangle as a Pedagogical Framework for Flipped-Class Instructional Videos in Introductory Chemistry. *J. Chem. Educ.* **2020**, *97* (6), 1536−1542.

(48) Brame, C. J. Effective Educational Videos: Principles and Guidelines for Maximizing Student Learning from Video Content. *CBE-Life Sci. Educ.* **2016**, *15* (4), No. es6.

(49) Urban, S.; Brkljača, R.; Cockman, R.; Rook, T. Contextualizing Learning Chemistry in First-Year Undergraduate Programs: Engaging Industry-Based Videos with Real-Time Quizzing. *J. Chem. Educ.* **2017**, *94* (7), 873−878.

(50) Flower, D.; James, M. L. *Small Teaching Online: Applying Learning Science in Online Classes*, 1st ed.; Jossey-Bass: San Francisco, CA, 2019.

(51) Atieh, E. L.; York, D. M. Through the Looking CLASS: When Peer Leader Attitudes Are Not What They Seem. *J. Chem. Educ.* **2020**, *97* (8), 2078−2090.