Selection Process of Students for a Novel STEM Summer Bridge Program

Dr. Margaret E. Beier, Rice University

Margaret Beier is an Associate Professor of Psychology at Rice University in Houston, TX. She received her B.A. from Colby College, and her M.S. and Ph.D. degrees from the Georgia Institute of Technology. Margaret's research examines the predictors of performance in educational and occupational settings. In particular, she is interested in the effects of examining gender, age, ability, personality, motivation, and self-regulation on a range of outcomes. She is a member of the American Educational Research Association and a Fellow of the Society for Industrial and Organizational Psychologists.

Dr. Ann Saterbak, Rice University

Ann Saterbak is a Professor in the Practice in the Biomedical Engineering Department at Duke University. She taught at Rice University from 1999 to 2017. At Rice, Saterbak introduced problem-based learning in the School of Engineering and more recently launched a successful first-year engineering design course taught in the Oshman Engineering Design Kitchen. Saterbak is the lead author of the textbook, Bioengineering Fundamentals. Saterbak's outstanding teaching was recognized at Rice through university-wide and departmental teaching awards. In 2013, Saterbak received the ASEE Biomedical Engineering Division Theo C. Pilkington Outstanding Educator Award. For her contribution to education within biomedical engineering, she was elected Fellow in the Biomedical Engineering Society and the American Society of Engineering Education.

Ms. Megan Rose McSpedon,

Megan McSpedon is the Associate Director of the Rice Emerging Scholars Program. She has been with the program since it was founded in 2012. Megan received a B.A. in English from Rice University.

Prof. Michael Wolf, Rice University

Michael Wolf is Professor of Mathematics at Rice University as well as Faculty Director of the Rice Emerging Scholars Program, an initiative he co-founded in 2012. The Rice Emerging Scholars program is a collection of interventions, beginning the summer before matriculation and extending for at least 2 years and often more, for a group of matriculating Rice STEM students whose preparation for STEM is weaker than those of their peers. All identified obstructions to success are addressed, beginning with a bridge program built around the most difficult scientific topics the students will meet as freshmen.

Selection Process of Students for a Novel STEM Summer Bridge Program

Abstract

This NSF Grantee Poster explores the selection process for Rice University's Emerging Scholar Program (RESP). Developed in June 2012, RESP is a comprehensive summer bridge and term-time advising program aimed at increasing STEM retention, graduation, and achievement in promising students who attended under-resourced high schools. RESP is not a remedial program, nor even an 'early college course' program. Rather, RESP aims to target deficits in K-12 preparation that may create undue obstructions for the program's participants (named Scholars in the program and this paper) compared to their peers. The objective of the non-credit summer bridge portion of the program is to prepare Scholars for the pace, rigor, and depth of the STEM curriculum at Rice University. This is achieved through exposure to the most challenging portions of freshman calculus, chemistry, and physics with special focus on complex word problems. During subsequent years, Scholars receive intensive and intrusive term-time advising from staff devoted to the program.

RESP Scholars are admitted to Rice through the regular admissions process. After accepting a spot in the entering class, these students are invited to attend the bridge program in the summer before their freshman year. Scholar admittance occurs independent of consideration for, or participation in, RESP. Scholars are selected through partnerships with Rice's Office of Admissions and other groups on campus. RESP partners with the Office of Admissions to review student admission information including SAT/ACT test scores, SAT subject test scores, first-generation status, academic ambitions and high school competitiveness ranking.

A separate principal selection mechanism for RESP is a novel diagnostic exam created in conjunction with the Schools of Natural Sciences and Engineering. The 11-question exam covers conceptual knowledge and tests skills in mathematics, chemistry, and physics with quantitative word problems that students are expected to know prior to arrival at Rice University. By focusing on applied problems and conceptual knowledge, the exam demonstrates a student's academic preparation, not their intellectual ability.

The current study examines the validity of the RESP diagnostic exam and its predictive validity relative to standardized tests with a sample of students ($N = 976$) who matriculated into Rice University from 2012 to 2014. The RESP diagnostic exam was related to grades, and we found that the correlation between the RESP diagnostic exam and grades was greater for STEM grades than non-STEM grades. We found that the diagnostic exam accounted for an incremental 9% of variance in STEM grades above SAT performance, but only 1% of incremental variance above SAT in non-STEM grades. Moreover, we found evidence of range restriction for both SAT and RESP diagnostic exam performance for Rice University matriculants, further suggesting the utility of the diagnostic exam is at the lower end of the distribution. In summary, our results suggest that an additional diagnostic exam written by schools to specifically measure STEM preparation for their program can be a useful addition to procedures for selecting students for special experiences such as summer bridge programs.

Introduction

Predicting success in post-secondary environments has been a focus of educational psychologists and administrators for decades [1]. Traditional measures like high school GPA and standardized test performance account for some variance in college performance and remain the most common predictors used in post-secondary environments. Nonetheless, questions remain about how we can improve upon these traditional measures to better select students and ensure student success [2].

The validity of selection procedures becomes even more important when one considers the prevalence and cost of attrition from universities. According to the National Center for Educational Statistics, more than one in four students at Title IV institutions (Pell Grant participating institutions) fails to complete a degree within 6 years, representing a major concern for student, parents, and the institutions themselves [3]. This level of failure is particularly troubling in light of the cost of post-secondary education. In the domains of Science, Technology, Engineering, and Mathematics (STEM), there is also concern that significant attrition can negatively impact American global competitiveness in these fields [4].

One way that institutions have addressed student attrition is to provide summer bridge programs for matriculating first-year students, particularly directly toward those who might be at risk for dropping out [5]. The Rice Emerging Scholar Program (RESP) is a comprehensive program designed to address attrition specific to STEM disciplines. In particular, this program selects students who are likely under-prepared for post-secondary experiences in STEM at Rice University.

With limited resources, it is important to select the most appropriate students for this program. Specifically, we are interested in evaluating the usefulness and validity of the selection procedures used to identify students who are offered positions in RESP. In addition to standard admission measures (e.g., SAT, SAT Math), the team has also developed and implemented a "home-grown" diagnostic test that specifically assesses preparation in multi-step, STEM word-problems. Thus, the purpose of this work is to validate the RESP diagnostic exam used for RESP selection and evaluate its utility for selecting appropriate students.

Rice University's Emerging Scholar Program

The Rice Emerging Scholar Program (RESP) is a comprehensive, STEM retention and achievement program aimed at Rice University undergraduates [6]. RESP comprises two parts: a six-week comprehensive summer bridge program occurring the summer before students matriculate at Rice, and ongoing term-time advising and support until a student graduates. Evidence suggests that the RESP program is successful in affecting student retention in college and in STEM [6].

Begun in June 2012, RESP targets admitted Rice undergraduates of high intellectual ability who attended under-resourced high schools, leading to weaker preparation and a higher rate of STEM attrition compared to peers. For example, some Scholars did not have access to calculus in high school, while many of their non-RESP peers completed rigorous AP BC calculus courses. This weak math background creates a challenge for many Scholars whose degree plans require calculus and calculus-based physics courses to be taken simultaneously in the first year.

As described elsewhere, RESP is not a remedial program [6]. Instead, the summer portion of RESP prepares Scholars for the pace, rigor, and depth of the STEM curriculum by focusing on the most challenging concepts covered in the first-year STEM curriculum. Taught almost entirely by Rice faculty, Scholars complete non-credit coursework in calculus, chemistry and physics. Scholars attend classes five days a week and complete two midterms and a final for each course. Scholars participate in daily study groups facilitated by trained upperclassmen, who also lead sessions focused on common college transition issues, acculturation, and STEM experience. Scholars also complete short modules in engineering design and problem solving.

During the academic year, RESP supports Scholars comprehensively and proactively [6]. This is achieved through an intrusive and intensive advising model in which all Scholars meet with professional program staff weekly or bi-weekly through their first two years at Rice. We term this advising model 'intrusive' given its proactive nature [7]. Staff provides academic advising as well as guidance on non-academic issues that may impact course performance. The goal of this proactive model is to identify and address issues early, before problems have permanent consequences. Additionally, RESP provides free tutoring, guidance on research opportunities, and ongoing programming to address student life issues.

Selection Process and RESP Diagnostic Exam

Students participating in RESP are admitted through the university's regular admissions process. After accepting a spot in the entering class, participants are selected through partnerships with the Office of Admissions and other programs. Considerations for inclusion in the RESP program include standardized test scores (SAT/ACT), SAT subject test scores, first-generation status, and high school competitiveness ranking. While not criteria in any way, most Scholars meet stringent definitions of low-income, and many are members of under-represented minority groups.

In addition to these factors, selection relies heavily on a diagnostic exam designed to assess knowledge of high school level STEM concepts and ability with multi-step word problems. Designed by Rice University faculty, this novel diagnostic exam was created in conjunction with faculty in the Schools of Natural Sciences and Engineering. The 11-question exam covers conceptual knowledge and tests skills in calculus, chemistry, and physics. The test contains only complex, multi-step, quantitative word problems. The questions refer to knowledge that students are expected to know prior to matriculation. In this way, the RESP diagnostic exam is explicitly designed to demonstrate a student's academic preparation, not their intellectual ability. See the Appendix for three example problems.

The exam is strongly encouraged, but voluntary, for STEM matriculants. The multiple-choice test is administered online through the university course management system. Students have access to the test shortly after admission to the university. With one-point assigned to each question, the average score on the exam is 6.9 of 11 points. Serious consideration for inclusion in the RESP program is given for students who earn six of 11 points or below.

Research Methods

Hypotheses. Several years into the program, the authors decided to conduct an analysis of its current Scholars selection strategy that uses a cluster of admissions data, plus the RESP diagnostic test. Our overall hypothesis is that the RESP diagnostic test is an important predictive tool for identifying students who are most at risk of persistence in STEM. From this we identified two specific hypotheses:

Hypothesis 1: The RESP diagnostic test is a valid test, and it correlates well with other measures. We hypothesize that the RESP diagnostic test correlates well with other standardized test, such as SAT Math. Given the math content of the RESP diagnostic exam, we expect it to be more highly correlated with STEM course grades than with grades outside of STEM.

Hypothesis 2: Because the RESP diagnostic test assesses the types of skills necessary for success in STEM courses (i.e., STEM preparedness), it gives additional, valuable information beyond the nationally normed tests, such as SAT, which are more strongly linked with ability. Specifically, we predict that the RESP diagnostic test accounts for incremental variance in STEM course performance above the SAT and SAT Math.

Participants. Data from the Office of Admissions and Office of the Registrar at Rice University were used for matriculation years 2012 – 2014. These are the only years in which both RESP diagnostic exam and grades are available. There were 1,787 matriculating students included in the initial sample; data for the SAT were available for 1,588 of them. Of this group, 976 had taken the voluntary diagnostic exam and thus compose the sample for this study. The study complies with university-approved IRB procedures.

Predictor variables. Admission data included each student's high scores on the SAT Total (1,600 points available) and SAT Math (800 points available). Data from the RESP diagnostic exam was the third predictive variable. The 11-item RESP diagnostic measure has an acceptable internal reliability ($\alpha = 0.76$) given the breadth and depth of items covered in the assessment [8].

Outcome variables. Average grades were calculated using course data provided by the Office of the Registrar. We converted student grades for each course into a numeric scale (A+ = 13, A = 12, A- = 11, … F = 1). We identified STEM and non-STEM courses from student records using the Classification of Instructional Programs (CIP) Codes developed by the U.S. Department of Education [9]. For each student, the average grades for STEM courses and non-STEM courses were then calculated.

Statistical Methods. We assessed the validity of the diagnostic exam using the approach recommended by Urbina [10]. Evidence for construct validity is gathered by examining convergent and discriminant relations between the new and established measures; performance on a new measure is expected to be more highly correlated with existing measures of similar constructs and not as highly correlated with existing measures of different constructs. Evidence for criterion-related validity is gathered through examining the relationship between any new measure and an external outcome. In the current study, the SAT represented a similar construct, and the external outcomes were grades in STEM and non-STEM courses.

Although examining correlations provides some information about how performance on assessments co-vary, regression analyses allows examination of the predictive validity of a single measure while controlling for other measures [11]. In the current study we were interested in whether the diagnostic exam was useful for predicting grades over and above standard measures such as the SAT. Thus, we controlled for SAT performance in a regression analysis to assess the incremental prediction of the diagnostic test.

Results

First, we compared the SAT performance of students who did and did not take the RESP diagnostic exam. There was a significant mean difference in SAT performance favoring students who took the diagnostic exam (mean $\pm$ standard deviation = 1480 $\pm$ 94) compared to those who did not (1451 $\pm$ 131) (t-test, $t(411.17) = 3.89$, $p < 0.01$). However, since the remainder of the study involves only students who did take the test, this difference does not impact the work.

Table 1 shows the descriptive statistics and inter-correlations for the five study variables. The predictive variables are SAT Total performance, SAT Math performance, and the RESP diagnostic test; the outcomes are Rice University grades in STEM and non-STEM courses. The correlation between SAT Math performance and the diagnostic exam was significant and medium in magnitude ($r = 0.491$, $p < 0.001$) (Table 1). The correlation between SAT Total performance and the diagnostic exam was lower (r=0.365, $p < 0.01$). This was expected, since the diagnostic test includes and relies on math knowledge and skills.

Table 1. Descriptive statistics of the five variable of interest (mean $\pm$ standard deviation). Also shown are inter-correlations of predictors (SAT Total, SAT Math, RESP diagnostic; marked in white) and outcomes (STEM grade, Non-STEM grade; marked in gray). All correlations are significant at the $p < 0.001$ level (N = 976).

Variable	Mean (SD)	1. SAT Total	2. SAT Math	3. Diagnostic	4. STEM grade	5. Non-STEM grade
1. SAT Total	1475 (103)	1.000				
2. SAT Math	747 (57)	0.711	1.000			
3. Diagnostic	6.87 (2.44)	0.365	0.491	1.000		
4. STEM Grade	9.98 (1.82)	0.408	0.435	0.441	1.000	
5. Non-STEM Grade	11.1 (1.33)	0.360	0.294	0.265	0.625	1.000

Providing support for its validity, the RESP diagnostic exam was correlated with STEM grade ($r = 0.441$, $p < 0.001$) (Table 1). Note that this correlation was greater in magnitude than the correlation between the diagnostic exam and grades for non-STEM courses ($r = 0.265$, $p < 0.001$). The difference in correlations between the STEM and non-STEM grades was significant ($z = 6.89$, $p < 0.001$) [12].

Table 1 also shows that the correlation between STEM grades and SAT Math ($r = 0.435$) is similar in magnitude to the correlation between STEM grades and the RESP diagnostic exam ($r = 0.441$). Given the relatively high correlation between SAT Math and the diagnostic ($r = 0.491$), this calls into question whether the diagnostic exam is a useful tool for predicting performance over and above the SAT Math. That is, if the SAT Math accounted for the same variance in performance as the RESP diagnostic exam, the diagnostic exam would be redundant with the SAT Math and there would be no value in using the diagnostic exam.

A stronger test of the validity of the diagnostic exam for predicting performance in STEM would examine whether it accounted for significant variability in grades after controlling for SAT performance. To this end, hierarchical regressions were conducted using STEM grades and non-STEM grades as outcomes. SAT Total score was entered as the first step. The RESP diagnostic exam performance was entered as the second step.

In Table 2, results from the regression analysis are shown in the first four columns for STEM grades and in the last four columns for non-STEM grades. In the table, R^2 is the amount of variance in the outcome (STEM or Non-STEM grades) accounted for by the predictors in the regression equation. ΔR^2 is the incremental variance accounted for when the diagnostic exam is included in the regression equation. For STEM grades, SAT Total performance accounted for 17% of the variance (R^2). When the diagnostic exam was entered into the model in step 2, it accounted for an incremental 9% of variance (ΔR^2) in STEM grades.

The Model 2 regression equation for STEM grades is $Y' = 0.287 + 0.006X_{1(\text{SAT Total})} + 0.235X_{2(\text{diagnostic})}$. The magnitude of the incremental variance (9%) is significant and large in magnitude relative to other assessments designed to predict post-secondary performance [2]. It is particularly challenging to account for variance in grades over and above SAT performance.

Table 2. Regression coefficients for first and second model equations. B is the unstandardized beta-weight and SE(B) is the standard error of B. ** is p <0.001.

	STEM Grades				Non-STEM Grades			
	B	SE(B)	R^2	ΔR^2	B	SE(B)	R^2	ΔR^2
Model 1								
SAT Total	0.008	0.00	0.17**		0.004	0.000	0.10**	
Model 2								
SAT Total	0.006	0.000			0.004	0.000		
Diagnostic	0.235	0.019	0.27**	0.09**	0.063	0.015	0.11**	0.01**

For non-STEM grades, SAT Total accounted for 10% of the variance (R^2). When the diagnostic exam was entered in step 2, it accounted for an additional 1% of incremental variance (ΔR^2) in non-STEM grades. The Model 2 regression equation for non-STEM grades is $Y' = 5.180 + 0.004X_{1(\text{SAT Total})} + 0.063X_{2(\text{diagnostic})}$. Thus, the diagnostic accounted for considerably more variance in STEM grades than non-STEM grades.

We also examined the incremental variance accounted for by the RESP diagnostic exam over performance on the SAT Math performance. We found that the diagnostic accounted for 7% incremental variance (ΔR^2) over SAT Math performance for STEM grades, and 2% of incremental variance (ΔR^2) over SAT Math performance for grades in non-STEM courses, again supporting the predictive validity of the RESP diagnostic exam. In this case, the magnitude of the incremental variance over and above the SAT Math is quite high (7%) – especially given the correlation between the RESP diagnostic exam and SAT Math (Table 1) and the difficulty of deriving incremental prediction above the SAT [2].

Because RESP typically selects students who score three or more points below the average of the diagnostic exam, we were interested in examining the predictive validity of the diagnostic exam in the lower range of SAT Math performance. We examined the scatterplot of student scores of the RESP diagnostic and the SAT Math (Figure 1). Each point or circle on the scatterplot represents the number of prospective students who earned particular scores on diagnostic exam and SAT Math. The scatterplot shows a restriction of range in performance at the highest levels of the distribution, as evidenced by the high number of scores in the upper right side of the figure. In other words, a large number of incoming students score at the top end of the SAT Math, the diagnostic the exam, or both. This restriction of range attenuates the value of both the diagnostic exam and the SAT Math exam for selecting students in to the university more generally. Nonetheless, there are definitely students who score 6 or less on the RESP diagnostic exam, and these students receive serious consideration for the program. Thus, the variability at the lower end of the distribution appears to lend additional evidence that the RESP diagnostic exam is a valuable tool for selecting students into RESP at the lower end of the SAT Math distribution.

Discussion and Future Direction

This study has focused on evaluating the utility of a "home-grown" STEM-focused diagnostic test that is applied at Rice University and is used a criteria for admission in the RESP program. Our results are based on three years of data on the diagnostic exam and course performance.

In addressing our first hypothesis, this study provided evidence for the predictive validity of an 11-item diagnostic exam for selecting students as RESP Scholars. That is, the RESP diagnostic test was significantly correlated with STEM grades, and this correlation was greater in magnitude than the correlation between the diagnostic and non-STEM grades.

In addressing our second hypothesis, our results suggest that the RESP diagnostic test is indeed predictive of performance, particularly in STEM grades. The diagnostic exam accounted for independent variance in grades over and above SAT Total performance and SAT Math performance. The RESP diagnostic may be most useful at the lower end of the distribution, although this is an idea that needs further investigation.

In summary, our results suggest that an additional diagnostic exam written by a school to specifically measure STEM preparation for its program can be a useful addition to selection procedures when nominating students for special experiences such as summer bridge programs. Preparation of such an exam needs to be done in conjunction with STEM faculty, including those who teach first-year, required STEM courses.

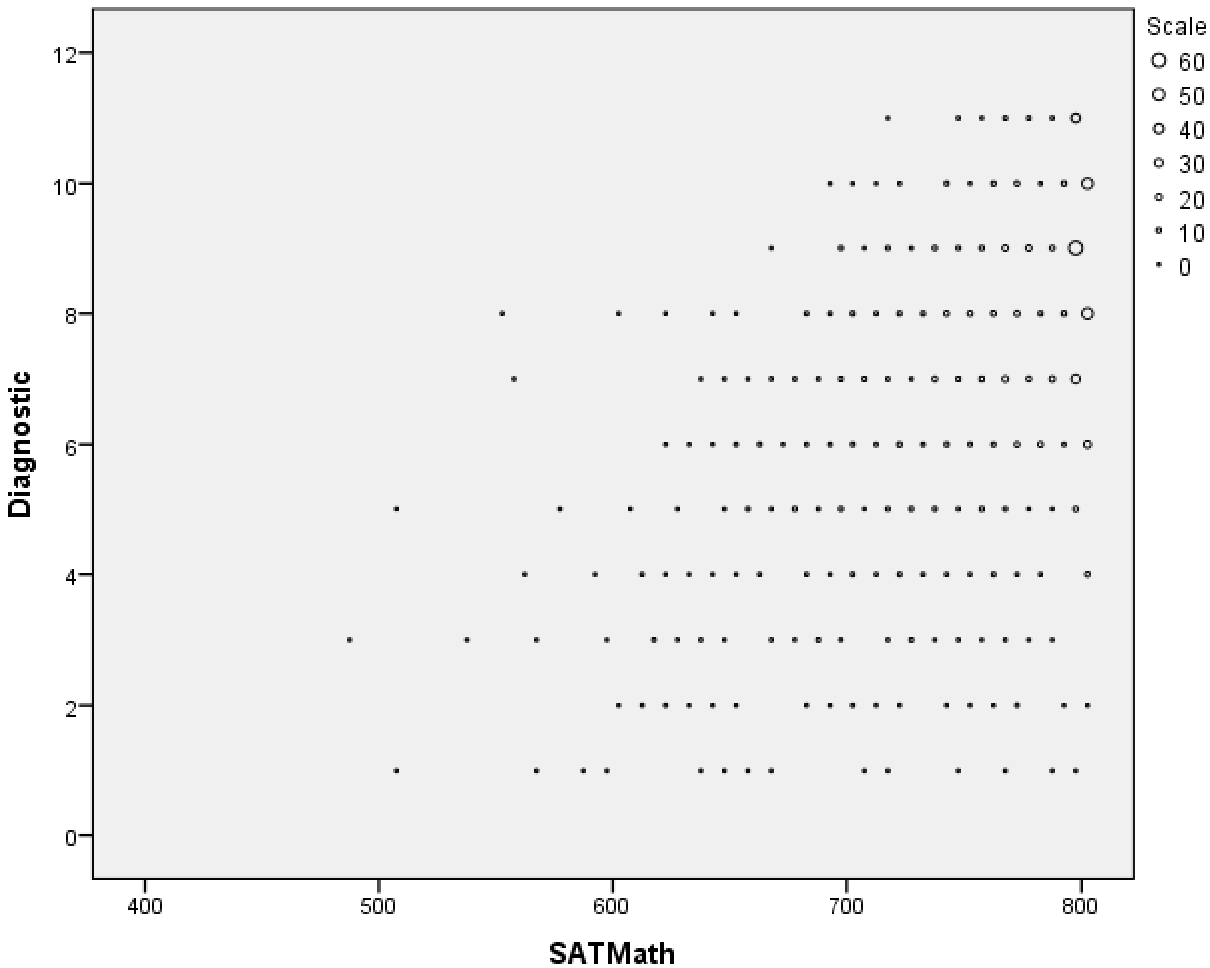


Figure 1. Scatterplot of individual scores on the RESP diagnostic exam and SAT Math performance. Scale legend shows graphics used for points with 0 – 9 through 60 people per scores on SAT Math and the Diagnostic test.

This study highlights the importance of targeting the selection of students into bridge programs, and other educational interventions. Specifically, our results suggest that the diagnostic test used here can be a useful determinant of performance. We will focus next on understanding the elements of the diagnostic exam (e.g., content/items) that are most useful for predicting STEM performance and then refining the measure. Other next steps include examining predictive validity of the RESP diagnostic exam for an array of outcomes including success within the RESP program itself.

Acknowledgements

This RESP program is partially supported by an NSF S-STEM program grant (#1565023). Other significant funding comes from Rice University. The research component of this program is partially funded by the S-STEM grant, and partially funded by Rice University and the Chao Foundation.

References

[1] Ackerman, P. L., Kanfer, R., & Beier, M. E. (2013). Trait complex, cognitive ability, and domain knowledge predictors of baccalaureate success, STEM persistence, and gender differences. *Journal of Educational Psychology*, *105*(3), 911–927. https://doi.org/10.1037/a0032338

[2] Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. Psychological Bulletin, 138(2), 353–387. https://doi.org/10.1037/a0026838

[3] National Center for Educational Statistics (2012). CES datalab tables. Retrieved from www.nces.ed.gov/datalab/tableslibrary/viewtable.aspxtableid7506

[4] Xie, Y., & Killewald, A. A. (2012). *Is American science in decline?* Cambridge, MA: Harvard University Press. doi:10.4159/harvard .9780674065048

[5] Contreras, F. (2011). Strengthening the bridge to higher education for academically promising underrepresented students. *Journal of Advanced Academics*, *22*(3), 500–526.

[6] McSpedon, M., A. Saterbak, M. Wolf. "Summer Bridge Program Structured to Cover Most Demanding STEM Topics." American Society of Engineering Education Annual Meeting, New Orleans, LA, June 2016.

[7] Kuh, G. D., (2007) "How to Help Students Achieve." Chronicle of Higher Education http://www.chronicle.com/article/How-to-Help-Students-Achieve/31980, accessed 27 March 2017.

[8] Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334.

[9] U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. Retrieved from: https://nces.ed.gov/ipeds/cipcode/default.aspx?y=55 on 12 Feb 2017.

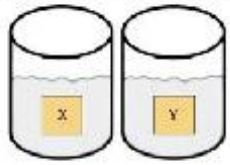
[10] Urbana, S. (2014). Essentials of psychological testing (2nd ed). Hoboken, NJ: Wiley.

[11] Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Applied multiple regression/correlation analysis for the behavioral sciences (3rd ed.). Mahwah, NJ: Earlbaum.

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

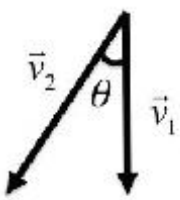
Appendix

4: As shown below, 100 g block of substance X has been heated to 100°C and dropped into a beaker of 100 ml of water at 25°C. A 100 g block of substance Y has also been heated to 100°C and dropped into another beaker of 100 ml of water at 25°C. The final temperature of the water in the beaker containing substance X is greater than that of substance Y. Which substance has the higher specific heat capacity (per gram)?



- A) Substance X
- B) Substance Y
- C) The heat capacities are equal
- D) Cannot be determined from the information given
- E) I don't know how to approach or answer this question.

5: At a given instant a stationary observer on the ground sees raindrops falling vertically with speed v_1. A passenger riding in a car that is traveling horizontally sees the raindrops falling with speed v_2 at some angle θ to the vertical. The speed of the car relative to the ground is



- A) $\sqrt{v_1^2 + v_2^2}$
- B) $\sqrt{v_2^2 - v_1^2}$
- C) $\sqrt{v_1^2 - v_2^2}$
- D) $v_2 - v_1$
- E) $v_1 + v_2$
- F) I have no idea how to answer this question.

6: A stalled car is being pushed off the road by a police cruiser as shown. If the car is being accelerated, we can conclude that:

- A) Because the car's engine is not running and its brakes are off, it cannot exert a force on the police cruiser.
- B) Because the presence of the car in front of the police cruiser reduces its acceleration, the size of the force exerted by the car on the cruiser must be greater than the size of the force exerted by the cruiser on the car.
- C) The net resultant force on both the car and the police cruiser is zero.
- D) The size of the force exerted by the police cruiser on the car equals the size of the force exerted by the car on the police cruiser.
- E) Because the car is accelerated, the size of the force exerted by the police cruiser on the car must be greater than the size of the force exerted by the car on the police cruiser.
- F) I have no idea of how to answer this question.