

Rural and First Generation Performance Differences on the Force and Motion Conceptual Evaluation

Rachel Henderson,¹ Cabot Zabriskie,² and John Stewart²

¹Department of Physics and Astronomy, Michigan State University, 567 Wilson Rd., East Lansing, MI, 48824

²Department of Physics and Astronomy, West Virginia University, 135 Willey St, Morgantown, WV, 26506

Differences in student performance on physics conceptual inventories have been studied with respect to gender and race/ethnicity. The current study expands this literature by exploring differences between first generation college students and rural/non-rural students on the Force and Motion Conceptual Evaluation (FMCE) using a large sample ($N = 3325$). Hierarchical linear regression was used to explore the effects of gender, race/ethnicity, first-generation status, and rural status. Significant differences in FMCE post-test scores were found by gender (14%), race/ethnicity (7%), first generation status (4%), and rural status (5%). Prior preparation, measured by ACT/SAT math scores, explained much of the performances by race/ethnicity and first generation status, but did not explain the differences in post-test scores by gender or rural status. No significant interactions between the different demographic features were measured.

I. INTRODUCTION

In many physics courses, students' conceptual understanding of Newton's laws of motion is measured using either the Force Concept Inventory (FCI) [1] or the Force and Motion Conceptual Evaluation (FMCE) [2]. Madsen, McKagan, and Sayre summarized the literature investigating the "gender gap" where, on average, male students outperform female students by 12% on mechanics conceptual inventories [3]. Many studies have explored the factors that may contribute to the gender gap such as prior preparation [4, 5], sociocultural factors [6], and item-level bias [7].

While the differences in conceptual understanding between male and female students have been extensively studied, in 2016, Scherr called for the Physics Education Research (PER) community to explore differences between other underrepresented populations within the physics classroom [8]. Hazari, Tai, and Sadler identified differences in physics grades between students of different races and ethnicities [9]. Recent research has reported differences in conceptual post-test scores by race and ethnicity for multiple PER instruments [10, 11]; these differences were independent of differences by gender and no gender-by-race/ethnicity interaction was measured. Prior academic preparation strongly mediated the effect of race/ethnicity, but not the effect of gender, on post-test scores.

First generation college students (FGCS) are students for whom neither parent nor guardian attended post-secondary school. Many studies have found that FGCS are less likely to enroll in 4-year colleges, attend less prestigious schools despite achievement, and persist at consistently lower rates than their non-first-generation peers [12–14]. Similar results are found for FGCS retention within Science, Technology, Engineering, and Mathematics (STEM) programs [15, 16]. FGCS are often from lower socioeconomic status (SES) households than non-first-generation students and cite financial reasons for departure from college at a higher rate [14]. Post-degree FGCS find similar career success to non-FGCS in employment and salary, but enroll in graduate or professional educa-

tion at a lower rate [12, 13].

A multinational study comparing mathematics achievement for rural and urban students found few differences between these two areas of the US [17]; instead, students' SES had the greatest impact on mathematics achievement. In 2017, a study by the College Board showed that students from rural areas have less access to Advanced Placement (AP) courses than students from urban or suburban areas; however, a similar percentage of rural students who took an AP test passed it with a score a 3 or better. For those students who did score a 3 or better on the AP test, there were similar enrollment and retention rates in college [18]. A study of rural and urban students in Oregon found that both groups enrolled in college immediately following high school at the same rate, but rural students persisted at lower rates. Controlling for achievement did not completely control for the difference in persistence rate, though higher achieving students did persist longer than lower achieving rural students [19].

Little research has investigated FGCS and rural students' performance in physics courses. This study will compare FMCE post-test performance for FGCS and rural students; these results will then be compared with those of other demographic groups. This study will investigate four demographic groups which are substantially in the minority in the physics classes at the institution studied (and in most institutions in the US). This study will address the following research questions: *RQ1: How is FGCS status and rural status related to conceptual physics knowledge? How does this compare with the relation of race/ethnicity and gender to conceptual knowledge? Does the interaction of these features further affect conceptual knowledge?* *RQ2: How is the effect of FGCS status, rural status, race/ethnicity, and gender on conceptual knowledge related to prior academic preparation?*

II. METHODS

This research was conducted between the spring 2011 and spring 2017 semesters at an eastern land-grant university

serving approximately 30,000 students. The general undergraduate population of the institution had a range of ACT scores from the 25th percentile to the 75th percentile of 21 to 26 and undergraduate demographic composition of 79% White, 6% international, 5% African American, 4% Hispanic, 2% Asian with other groups each 4% or less. Students in this study were enrolled in the introductory, calculus-based mechanics course. From fall 2011 to spring 2015, students attended 4 50-minute lectures and 1 two-hour laboratory per week. From fall 2015 to spring 2017, the course was modified to 3 50-minute lectures and 1 three-hour laboratory per week. The results for each time periods were statistically similar and, therefore, the data was aggregated. A Learning Assistants (LA) program was implemented for the period studied with LAs presenting research-based materials in the laboratory. The FMCE was administered pre- and post-instruction; credit for completion was awarded to the students for a good faith effort. A total of 5289 students completed the course for a grade. Students that received either a “D” or “F” international students, students with missing ACT and SAT scores, and students who did not complete both the FMCE pretest or post-test were eliminated leaving a total sample size of $N = 3325$ students. Students receiving a “D” or “F” were eliminated because very few students received these grades and the research team felt they should be treated as a different population.

Students’ demographic information including ACT/SAT scores, gender, race/ethnicity, first generation status and permanent address were collected from university records. ACT and SAT mathematics scores were converted to a percentile score using scales published by the testing companies; this percentile score will be called ACTM%. The following codes were used for all analyses: gender was coded with male as 0 and female as 1, race/ethnicity was coded with White non-Hispanic students as 0 and other students as 1 (non-White), and first generation status was coded with non-FGCS as 0 and FGCS as 1. While more fine-grained racial and ethnic data was available, disaggregating further would not have allowed the study of rural or first generation status.

The zip codes from the student’s permanent address were coded using the 2013 US Department of Agriculture (USDA) Rural-Urban Continuum Codes [20]. The USDA codes classify counties as either metropolitan or non-metropolitan based upon population size, degree of urbanization and adjacency to other areas using a 9-point score. Students were classified as rural using the last 3 classifications: “Urban population of 2500 to 19,999, not adjacent to a metro area,” “Completely rural or less than 2500 urban population, adjacent to a metro area,” and “Completely rural or less than 2500 urban population, not adjacent to a metro area.” Rural was coded as 1, non-rural as 0.

III. RESULTS

Table I presents the FMCE pretest and post-test percentages and ACTM% disaggregated by gender, race/ethnicity, FGCS status, and rural status. Differences between groups were analyzed using t -tests and effect sizes were characterized by Cohen’s d : $d = 0.2$ is a small effect, $d = 0.5$ is a medium effect, and $d = 0.8$ is a large effect.

TABLE I. FMCE post-test percentages and ACT/SAT math percentiles. The effect size d of the difference between groups is reported. Superscript “a” denotes $p < 0.05$, “b” denotes $p < 0.01$, and “c” denotes $p < 0.001$.

	Pretest%		Post-test%		ACTM%	
	$M \pm SE$	d	$M \pm SE$	d	$M \pm SE$	d
Male (2626)	25 ± 0.39	.35 ^c	54 ± 0.56	.48 ^c	80 ± 0.29	.02
Female (699)	19 ± 0.51		40 ± 0.93		80 ± 0.55	
White (2965)	24 ± 0.35	.19 ^c	52 ± 0.53	.24 ^c	81 ± 0.26	.32 ^c
non-White (360)	21 ± 0.83		45 ± 1.4		76 ± 0.98	
FGCS (455)	21 ± 0.81	.15 ^b	48 ± 1.3	.12 ^a	77 ± 0.81	.24 ^c
non-FGCS (2870)	24 ± 0.36		52 ± 0.53		80 ± 0.27	
non-Rural (3173)	24 ± 0.34	.20 ^b	51 ± 0.51	.17 ^a	80 ± 0.26	.08
Rural (152)	20 ± 1.3		47 ± 2.2		81 ± 1.2	

There were significant differences in the FMCE pretest [$t(1598) = 10.16, p < 0.001, d = 0.35$] and post-test [$t(1255) = 12.43, p < 0.001, d = 0.48$] scores between men and women; however, there was no significant difference in ACTM%. Men outperformed women by 6% on the pretest and by 14% on the post-test.

Significant differences between White students and non-White students were measured on the FMCE pretest [$t(463) = 4.54, p < 0.001, d = 0.19$] and FMCE post-test [$t(499) = 3.89, p < 0.001, d = 0.24$]; there was also a difference in ACTM% [$t(411) = 4.66, p < 0.001, d = 0.32$]. White students outperformed non-White students on all three measures, by 3% on the FMCE pretest, 7% on the FMCE post-test, and 5% on ACTM%.

Significant differences between FGCS and non-FGCS were also identified for the FMCE pretest [$t(645) = 3.11, p = 0.002, d = 0.15$], FMCE post-test [$t(615) = 2.43, p = 0.02, d = 0.12$], and ACTM% [$t(558) = 4.20, p < 0.001, d = 0.24$]. FGCS scored 3% lower on the pretest, 3% lower ACTM%, and 4% lower on the post-test.

Differences were also measured for rural and non-rural students: FMCE pretest [$t(173) = 2.86, p = 0.005, d = 0.20$] and post-test [$t(168) = 2.11, p = 0.04, d = 0.17$]. Rural students scored 4% lower on both the pretest and post-test; however, there was no difference in ACTM%.

To further explore the differences in FMCE post-test scores, as well as the interactions between membership in multiple demographic groups, hierarchical linear regression (HLR) was employed; results are presented in Table II and

TABLE II. Hierarchical linear regression predicting FMCE post-test percentage. Superscript “a” denotes $p < 0.05$, “b” denotes $p < 0.01$, and “c” denotes $p < 0.001$.

Model	Variables	B	SE	β	R^2	ΔR^2
1	Female	-13.55 ^c	1.2	-0.47 ^c	0.037 ^c	
2	non-White	-6.90 ^c	1.6	-0.24 ^c	0.006 ^b	
3	FGCS	-3.44 ^a	1.4	-0.12 ^a	0.002 ^a	
4	Rural	-4.73 ^a	2.4	-0.17 ^a	0.001 ^a	
5	Female non-White	-13.34 ^c -6.09 ^c	1.2 1.6	-0.47 ^c -0.21 ^c	0.042 ^c	0.005 ^c
6	Female FGCS	-13.67 ^c -4.02 ^b	1.2 1.4	-0.48 ^c -0.14 ^b	0.040 ^c	0.003 ^b
7	Female Rural	-13.59 ^c -5.10 ^a	1.2 2.3	-0.48 ^c -0.18 ^a	0.039 ^c	0.002 ^a
8	non-White FGCS	-6.86 ^c -3.37 ^a	1.6 1.4	-0.24 ^c -0.12 ^a	0.007 ^c	0.001 ^a
9	FGCS Rural	-3.28 ^a -4.41	1.4 2.4	-0.11 ^a -0.15	0.003 ^a	0.001

III. The tables present the regression coefficient B , the standard error SE, the normalized regression coefficient β , the R^2 for the model, and the change in ΔR^2 between the model and the simpler model in which it is nested. The significance of the coefficient, the model, or the improvement in model fit of the nested model is shown as a superscript. HLR is a statistical technique used to determine how much additional variance is explained in the dependent variable when adding one additional independent variable. The successive models are then compared using ANOVA to calculate the significance of the improved model fit. For all models, the FMCE post-test percentile score is the dependent variable. HLR differs from the similarly named Hierarchical Linear Modelling (HLM) which is used to investigate models with nested data.

Models 1-4 (Table II) present models with each demographic factor independently. The B coefficient measures the difference in post-test score for students in the demographic group as compared to students not in the group; the β coefficient measures the change in standard deviation units and may be used as an effect size (compare with Cohen’s d in the previous section). Women, non-White students, FGCS, and rural students had significantly ($p < 0.05$) lower average post-test scores than other students with women having the greatest difference in scores and FGCS having the least; however, the amount of variance explained (R^2) was fairly small for each group.

Models 5-9 (Table II) are nested within models 1-4 and examine combinations of demographic factors. The regression coefficients in these models changed little over the coefficients in the models in which they were nested (compare the gender coefficient in models 5-7 to the coefficient in model 1 for example). These models also explained little additional variance. Interactions between the demographic factors were

TABLE III. HLR analysis predicting FMCE post-test percentage controlling for ACTM%. Superscript “a” denotes $p < 0.05$, “b” denotes $p < 0.01$, and “c” denotes $p < 0.001$.

Model	Variables	B	SE	β	R^2	ΔR^2
0A	ACTM%	0.82 ^c	0.03	0.42 ^c	0.180 ^c	
1A	Female ACTM%	-13.85 ^c 0.82 ^c	1.1 0.03	-0.48 ^c 0.43 ^c	0.219 ^c	0.039 ^c
2A	non-White ACTM%	-3.04 ^a 0.81 ^c	1.5 0.03	-0.10 ^a 0.42 ^c	0.181 ^c	0.001 ^a
3A	FGCS ACTM%	-0.52 0.82 ^c	1.3 0.03	-0.012 0.42 ^c	0.180 ^c	0.000
4A	Rural ACTM%	-5.67 ^b 0.82 ^c	2.1 0.03	-0.19 ^b 0.42 ^c	0.181 ^c	0.001 ^b
5A	Female non-White ACTM%	-13.77 ^c -2.18 0.82 ^c	1.1 1.4 0.03	-0.48 ^c -0.08 0.42 ^c	0.219 ^c	0.000
6A	Female FGCS ACTM%	-13.88 ^c -1.10 0.82 ^c	1.1 1.3 0.03	-0.48 ^c -0.04 0.42 ^c	0.219 ^c	0.000
7A	Female Rural ACTM%	-13.89 ^c -6.05 ^b 0.82 ^c	1.1 2.1 0.03	-0.49 ^c -0.21 ^b 0.43 ^c	0.221 ^c	0.002 ^b
8A	non-White FGCS ACTM%	-3.04 ^a -0.51 0.82 ^c	1.5 1.3 0.03	-0.11 ^a -0.02 0.42 ^c	0.181 ^c	0.000
9A	FGCS Rural ACTM%	-0.31 -5.64 ^b 0.82 ^c	1.3 2.2 0.03	-0.01 -0.20 ^b 0.42 ^c	0.181 ^c	0.001 ^b

also tested; no interaction was significant. The combination of non-White and rural could not be investigated because only 5 students shared this combination; all other combinations contained at least 28 students.

Model 0A (Table III) shows the overall effect of ACTM% on post-test scores. Models 1A-4A control models 1-4 for ACTM%; the effect of controlling for ACTM% varied strongly by demographic group. ACTM% strongly mediated the effect of race/ethnicity reducing the regression coefficient by 50% (compare models 2 and 2A) and the effect of FGCS status (models 3 and 3A) reducing the regression coefficient until it was no longer significant. The effect of controlling for ACTM% on gender and rural status was completely different; the regression coefficients of both women and rural students increased in magnitude slightly. Models 5A-9A examined combinations of demographic factors controlling for ACTM%; these models explained little additional variance with models 5A, 6A, and 8A not significantly improving the models in which they were nested. The addition of a second demographic factor also had little effect on the regression coefficient of the first factor. No interaction between demo-

graphic factors was significant.

While the interactions between demographic factors were not significant, there was a significant gender-by-ACTM% interaction [$\beta = -0.12$, $p = 0.002$] and a significant race/ethnicity-by-ACTM% interaction [$\beta = -0.14$, $p < 0.001$]. The FGCS-by-ACTM% and rural-by-ACTM% interactions were not significant. As such, ACTM% scores are less predictive of post-test scores for women and non-White students, but not FGCS and rural students.

IV. DISCUSSION AND CONCLUSION

RQ1: How is FGCS status and rural status related to conceptual physics knowledge? How does this compare with the relation of race/ethnicity and gender to conceptual knowledge? Does the interaction of these features further affect conceptual knowledge? Each demographic category studied was related to significantly lower FMCE post-test scores. Gender and race/ethnicity were related to the largest differences (14% and 7% respectively). FGCS and rural status were related to smaller but significant differences (4% and 5% respectively). These differences were independent; no interaction between demographic factors was measured. Introductory physics courses are a requirement for most STEM majors; therefore, this 4% gap in FGCS conceptual understanding could contribute to the lower rate of college persistence of FGCS [15, 16]; however, further research is needed to investigate this hypothesis.

RQ2: How is the effect of race/ethnicity, gender, FGCS status, and rural status on conceptual knowledge related to prior academic preparation? Academic prior preparation measured by ACTM% mediated the relation of gender (small

effect) and race/ethnicity (large effect) on conceptual post-test scores as has previously been reported [10, 11]. The mediating effect of ACTM% on FGCS status was similar to its mediating effect on race/ethnicity with the FGCS regression coefficient reduced sufficiently such that it was no longer significant. The mediating effect of ACTM% on rural status was similar to its mediating effect on gender with the rural regression coefficient increasing in magnitude. As such, ACTM% explains the differences in post-test performance by race/ethnicity and FGCS, but not gender or rural status. The origin of the performance difference by gender is an active area of research [3]; perhaps the similarity with rural status may open a new productive path to explore the gender gap. The behavior of rural status and gender when controlling for ACTM% was not identical; a significant gender-by-ACTM% interaction was measured, but the rural-by-ACTM% interaction was not significant.

Differences in FMCE post-test performance were identified between members and non-members of four demographic groups: women, students not identifying as White non-Hispanic, FGCS, and rural students. The differences were largest for women and smallest for FGCS. The disadvantages to members of each group were generally independent with additional disadvantage accumulating for each additional group in which the student was a member. No interactions between the effect of membership in multiple groups were detected. Controlling for academic prior preparation with ACT/SAT math scores mediated the effect of the demographic factors differently for each group; the effect of race/ethnicity and FGCS was greatly reduced, while the effect of gender and rural status was slightly enhanced. This work was supported by the National Science Foundation under Grant No. EPS-1003907.

-
- [1] D. Hestenes, M. Wells, and G. Swackhamer, *Phys. Teach.* **30**, 141 (1992).
 - [2] R. K. Thornton and D. R. Sokoloff, *Am. J. Phys.* **66**, 338 (1998).
 - [3] A. Madsen, S. McKagan, and E. Sayre, *Phys. Rev. Phys. Educ. Res.* **9**, 020121 (2013).
 - [4] L. Kost, S. Pollock, and N. Finkelstein, *Phys. Rev. Phys. Educ. Res.* **5**, 010101 (2009).
 - [5] L. Kost-Smith, S. Pollock, and N. Finkelstein, *Phys. Rev. Phys. Educ. Res.* **6**, 020112 (2010).
 - [6] A. Miyake, L. Kost-Smith, N. Finkelstein, S. Pollock, G. Cohen, and T. Ito, *Science* **330**, 1234 (2010).
 - [7] A. Traxler, R. Henderson, J. Stewart, G. Stewart, A. Papak, and R. Lindell, *Phys. Rev. Phys. Educ. Res.* **14**, 010103 (2018).
 - [8] R. Scherr, *Phys. Rev. Phys. Educ. Res.* **12**, 020003 (2016).
 - [9] Z. Hazari, R. Tai, and P. Sadler, *Sci. Educ.* **91**, 847 (2007).
 - [10] R. Henderson and J. Stewart, in *AIP conference proceedings* (AIP, 2017), pp. 172–175.
 - [11] C. Zabriskie, S. DeVore, R. Henderson, J. Stewart, P. Miller, G. Stewart, and G. Cochran, *The relation of race, ethnicity, and gender to physics conceptual inventory performance* (2018), submitted *Phys. Rev. Phys. Educ. Res.*
 - [12] E. Pascarella, C. Pierson, G. Wolniak, and P. Terenzini, *J. High. Educ.* **75**, 249 (2004).
 - [13] E. Cataldi, C. Bennett, and X. Chen, *First-generation students: College access, persistence, and postbachelor's outcomes.*, National Center For Education Statistics (NCES, 2018).
 - [14] J. Redford and K. Hoyer, *First-generation and continuing-generation college students: A comparison of high school and postsecondary experiences.*, (NCES, 2018).
 - [15] X. Chen, *STEM attrition: College students' paths into and out of STEM fields.*, (NCES, 2013).
 - [16] D. Verdin and A. Godwin, in *Frontiers in Education Conference (FIE), 2015 IEEE* (IEEE, 2015), pp. 1–8.
 - [17] J. Williams, *J. Res. Rural Educ.* **20**, 20 (2005).
 - [18] *Advanced Placement Access and Success: How do rural schools stack up?*, The College Board, New York, NY (2017).
 - [19] A. Pierson and H. Hanson, *Comparing postsecondary enrollment and persistence among rural and non-rural students in Oregon*, US Department of Education, (2015).
 - [20] *Rural-Urban Continuum Codes: Documentation*, US Department of Agriculture, (2013).