Summer 2021
Volume 12, Issue Supplemental 1

# Examining the Impact of the BUilding Infrastructure Leading to Diversity (BUILD) Initiative on Academic and Researcher Self-Efficacy among First Year Students

**Krystle Palma Cobian[1], Shujin Zhong[1], and Lourdes Guerrero[1]**

[1]University of California, Los Angeles, Los Angeles, CA, 90025

Email: cobiank@ucla.edu

Keywords: Self Efficacy, Science Self-efficacy, Research Training

## Abstract

**Introduction.** Stemming from Albert Bandura's work on self-efficacy (1977) and later work on social cognitive theory (Bandura, 1991), self-efficacy was developed to explain how people think, motivate themselves, and ultimately how they behave - including how long they will persist in the face of obstacles or challenging situations (Bandura & Adams, 1977).

When considering the impact of self-efficacy for students with scientific career aspirations, especially for underrepresented and disadvantaged groups in the biomedical sciences, the literature suggests that self-efficacy is both an outcome of efforts to support students interested in science, technology, engineering, and math (STEM) (Carpi et al., 2017) as well as a driver of future behaviors, such as increasing aspirations for a STEM career (Amelink et al., 2015). Also, important to note is that self-efficacy can differ by social identities such as gender, race, and socioeconomic status (MacPhee, Farro, & Canetto, 2013).

Research on self-efficacy has examined relationships between students' beliefs about their abilities and their resulting academic achievement (Chemers, Hu, & Garcia, 2001; Honicke & Broadbent, 2016) as well as their career options (Byars & Hackett, 1998; Lent, Brown, & Hackett, 2000). Self-efficacy must be domain-specific, meaning that an individual's perception of self-efficacy will vary across specific spheres of activities. For example, academic self-efficacy is one's conviction in being able to successfully perform a given academic-related task at a designated level (Schunk, 1991).

For this study, we examine science self-efficacy as a researcher in order to understand what pre-college characteristics and freshmen experiences might impact self-efficacy in the domain of science skills. A student who experiences high academic self-efficacy in an introductory STEM course, for example, may have a low self-efficacy perception of their skills in a research lab, which

 

requires different competencies compared to learning scientific theories and concepts in a classroom setting. Thus, while academic self-efficacy and science identity have been studied with respect to STEM outcomes, little is known about how one develops self-efficacy regarding science and research skills.

The BUILD programs are designed to explore the most effective ways to engage students from underrepresented backgrounds in biomedical research, helping them progress on the pathway to become potential future contributors to the NIH-funded research enterprise. Given the research on scientific self-efficacy and its connection to degree completion and career interests, we explore how involvement in the BUILD program relates to changes in students' academic self-efficacy after they participated in the BUILD scholar program during their first year in college.

**Research Question.** Does participation in the BUILD scholar program during freshman year impact students' scientific self-efficacy?

**Methods. Sample.** Data are derived from the Higher Education Research Institute's Freshman Survey (TFS) and the DPC Student Annual Follow-Up Survey (SAFS). The sample consists of students at 4 BUILD institutions that had first-year BUILD programming and thus enrolled students into BUILD prior to, or at the beginning of their first academic year of college. Incoming freshmen completed the TFS before the fall of their first year and the SAFS during the spring of their first year. We assume that students' action of taking both surveys is random, and the missingness in the datasets is also due to randomness.

**Matching.** The sample consists of three cohorts (2016-2018) of first-year BUILD students, each of whom was matched with 2 of their non-BUILD peers, due to the fact that there were a lot more students in the control group (108 students in the treated group vs. around 20,000 students in the control group). We firstly extracted students who were BUILD scholars and had records of taking both TFS and SAFS. We identified 108 students from the 4 BUILD primary sites, and employed a two-step matching procedure using exact matching to ensure that we include students from the targeted institutions and cohorts, and using propensity score matching for baseline covariates: gender (SEX), race/ethnicity (RACEGROUP), pell-grant status (Pell), first generation status (first), high school GPA (HSGPA), and years of mathematics courses taken during high school (YRSTDY2). The summary of balance of the pre-propensity score matching data is reported in Table 1.

 

Table 1: Summary of Balance for All Data (Balance Check)

| | Means Treated | Means Control | Std. Mean Diff. | Var. Ratio | eCDF Mean |
|---|---|---|---|---|---|
| distance | 0.201 | 0.105 | 0.758 | 1.936 | 0.249 |
| SEX | 1.694 | 1.679 | 0.033 | 0.920 | 0.010 |
| AIDTYPE3 | 1.435 | 1.344 | 0.183 | 1.098 | 0.046 |
| firstg | 0.296 | 0.274 | 0.048 | NA | 0.022 |
| HSGPA | 6.824 | 6.882 | -0.049 | 0.938 | 0.017 |
| YRSTUDY2 | 5.944 | 5.954 | -0.016 | 1.274 | 0.005 |
| AMAJOR1 | 0.861 | 0.694 | 0.483 | NA | 0.167 |
| AMAJOR2 | 0.102 | 0.057 | 0.147 | NA | 0.045 |
| AMAJOR3 | 0.037 | 0.249 | -1.121 | NA | 0.212 |
| RACEGROUP2 | 0.176 | 0.182 | -0.015 | NA | 0.006 |
| RACEGROUP3 | 0.306 | 0.088 | 0.473 | NA | 0.218 |
| RACEGROUP4 | 0.111 | 0.185 | -0.236 | NA | 0.074 |
| RACEGROUP5 | 0.167 | 0.367 | -0.538 | NA | 0.200 |
| RACEGROUP6 | 0.028 | 0.012 | 0.095 | NA | 0.016 |
| RACEGROUP7 | 0.213 | 0.166 | 0.115 | NA | 0.047 |
| SciSE | 54.165 | 53.118 | 0.133 | 0.646 | 0.035 |

For the propensity score matching, we identified 2 control units per 1 treated unit through nearest neighbor matching without replacement, estimated with logistic regression. After matching, we obtained a relatively balanced sample in the treated and the control groups (see Table 1 vs. Table 2, and Figure 2). The improvement of balance can be reflected from the reduced standardized mean differences and the reduced differences in the empirical cumulative density function between the treated and control groups.

Table 2: Summary of Balance for Matched Data

| | Means Treated | Means Control | Std. Mean Diff. | Var. Ratio | eCDF Mean | Std. Pair Dist. |
|---|---|---|---|---|---|---|
| distance | 0.201 | 0.193 | 0.064 | 1.212 | 0.005 | 0.067 |
| SEX | 1.694 | 1.648 | 0.100 | 0.865 | 0.022 | 0.980 |
| AIDTYPE3 | 1.435 | 1.468 | -0.065 | 0.992 | 0.016 | 0.567 |
| firstg | 0.296 | 0.310 | -0.030 | NA | 0.014 | 0.943 |
| HSGPA | 6.824 | 6.903 | -0.067 | 1.078 | 0.018 | 0.959 |
| YRSTUDY2 | 5.944 | 5.981 | -0.066 | 1.540 | 0.007 | 0.644 |
| AMAJOR1 | 0.861 | 0.870 | -0.027 | NA | 0.009 | 0.455 |
| AMAJOR2 | 0.102 | 0.102 | 0.000 | NA | 0.000 | 0.148 |
| AMAJOR3 | 0.037 | 0.028 | 0.049 | NA | 0.009 | 0.049 |
| RACEGROUP2 | 0.176 | 0.181 | -0.012 | NA | 0.005 | 0.401 |
| RACEGROUP3 | 0.306 | 0.269 | 0.080 | NA | 0.037 | 0.101 |
| RACEGROUP4 | 0.111 | 0.139 | -0.088 | NA | 0.028 | 0.383 |
| RACEGROUP5 | 0.167 | 0.176 | -0.025 | NA | 0.009 | 0.124 |
| RACEGROUP6 | 0.028 | 0.032 | -0.028 | NA | 0.005 | 0.254 |
| RACEGROUP7 | 0.213 | 0.204 | 0.023 | NA | 0.009 | 0.226 |
| SciSE | 54.165 | 53.605 | 0.071 | 0.688 | 0.036 | 1.050 |

 

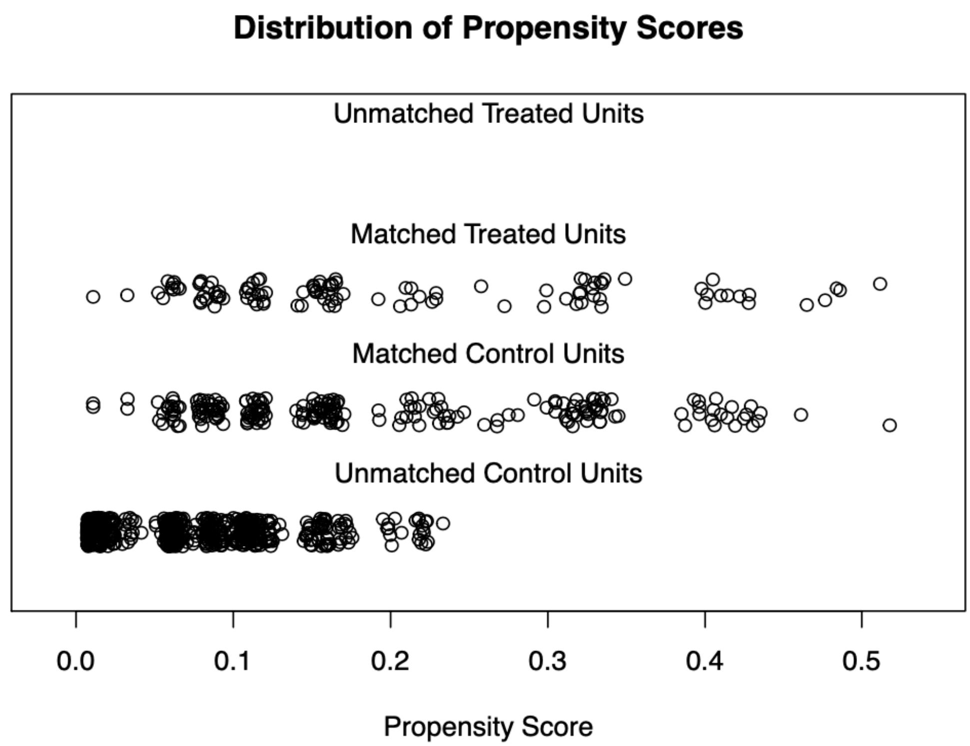


Figure 1: Comparisons

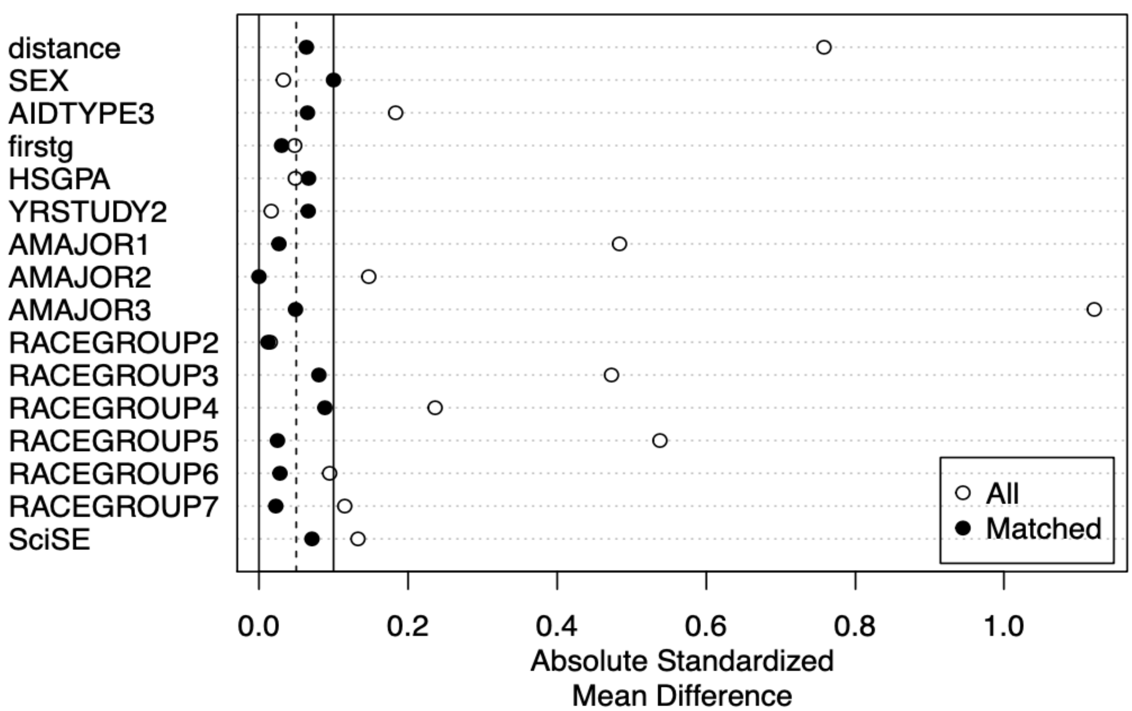


Figure 2: Absolute Standardized Mean Difference

 

**Regression Analysis.** We used the matched sample to perform regression analyses to see if, after controlling for covariates, the BUILD scholar program would be significantly influential in predicting science self-efficacy. We designed a time series model and defined students' science self-efficacy at time point t as our dependent variable.

The dependent variable is scientific self-efficacy as a researcher, conceptualized as the science self-efficacy construct on both the TFS and SAFS. Science self-efficacy is a measure of students' confidence in their ability to conduct scientific research. The original construct developed by the Higher Education Research Institute (HERI) consists of 10 items that asked participants about their perceptions to engage in various scientific skills. For this longitudinal study, the number of items in the construct was reduced to six items. The set of questions in the surveys asked students how confident they felt with the following science skills: using technical science skills, generating a research question, determining how to collect appropriate data, explaining the results of a study, using scientific literature to guide research, and integrating results from multiple studies.

The variable is quantified using students' expected-a-posterior (EAP) item response scores of six items in the surveys, and can be treated as a continuous variable. The scores are centered and scaled at N(50, 10). The intervention indicator is whether or not a student was in the BUILD scholar program between the time they took the TFS and SAFS. Based on related literature, covariates we plan to include in this study are gender, race/ethnicity, Pell-grant status, first-gen status, high school GPA, years of match training in high school, students' lab experience, conference participation, and faculty mentoring. Since there were multiple survey items related to faculty mentoring, we selected the following five items and computed EAP scares for each observation in the whole dataset: faculty showed concern about students' progress, faculty empower student to learn here, faculty believe in student's potential to succeed academically, faculty encouraged student to meet with them outside of class, and at least one faculty member has taken interest in student's development.

**Results.** We designed a time series model and defined students' scientific self-efficacy at time point t as our dependent variable. The basic model is similar to a pair-wised t-test, controlling for the grouping differences. Table 3 shows the results of the time series basic model. As time went by, students' scientific self-efficacy would decrease, the BUILD scholars' science self-efficacy tended to keep growing. The intra-class correlation (ICC) of the within person (case) random effect is 0.432, which indicates the necessity of including the random effect, or the variance contributed by individual differences (one's post-survey result only comparing with their own pre-survey measures).

 

**Time Series Basic Model (Pairewise t-test)**

$SciSE_t = \beta_0 + \beta_1 \times scholar + \beta_2 \times time + \delta_{pair} + \epsilon$, where $SciSE_t$ is the depedent variable at time point $t$.

Table 3: Time Series Basic Model (Pairewise t-test)

| | Estimate | Std. Error | t statistics | sig.level |
|---|---|---|---|---|
| (Intercept) | 0.379 | 0.051 | 7.487 | *** |
| scholar | 0.445 | 0.147 | 3.037 | ** |
| time | -0.440 | 0.081 | -5.412 | *** |
| Random Effects | Variance | Std. | | |
| case | 0.359 | 0.599 | | |
| Residual | 0.472 | 0.687 | | |
| ICC | 0.432 | | | |
| AIC | 1267.675 | | | |
| BIC | 1288.586 | | | |

Building on the basic model, we added students' background-related covariates: reported major (variable Major, coded as 0-non-biomedical related majors, 1-biomedical social science majors, and 2-biomedical natural science majors), high school GPA, gender, race/ethnicity, pell-grant eligibility, first generation college student status, and years of math training in high school. The results are reported in Table 4. The results show consistent effects of time and the BUILD scholar intervention. In addition, we observe that majors, especially students who are in the biomedical natural science majors (Major2 in Table 4), compared to non-biomedical majors tend to have higher science self-efficacy. Those who had higher GPAs in high school (HSGPA) also tended to have higher scientific self-efficacy. On the other hand, women (variable SEX), students who self-identified as Latina/x/o (RACEGROUP4), and who were first generation college students (firstg) appear to be less likely to have a high scientific self-efficacy at the end of their first year in college. The variance of the random effect reduced in this model, but the ICC remained around 40%, which was still large enough to be kept in the model.

 

### Time Series Model with Covariates

$SciSE_t = \beta_0 + \beta_1 \times scholar + \beta_2 \times time + \delta_{pair} + \boldsymbol{\beta} \cdot \mathbf{X_{cov}} + \epsilon$, where

$$\mathbf{X_{cov}} = (HighSchoolGPA, Math, Gender, Race/ethnicity, PellGrant, FirstGeneration)^T$$

The coefficients $\boldsymbol{\beta} = (\beta_3, \beta_4, ..., \beta_p)$, where $p$ is the number of variables to predict $SciSE_t$.

Table 4: Time Series Model with Covariates

| | Estimate | Std. Error | t statistics | sig.level |
|---|---|---|---|---|
| (Intercept) | 0.005 | 0.645 | 0.008 | |
| scholar | 0.443 | 0.143 | 3.095 | ** |
| time | -0.416 | 0.081 | -5.169 | *** |
| Major1 | 0.396 | 0.230 | 1.726 | . |
| Major2 | 0.587 | 0.195 | 3.014 | ** |
| HSGPA | 0.115 | 0.041 | 2.802 | ** |
| SEX | -0.220 | 0.097 | -2.256 | * |
| RACEGROUP2 | -0.032 | 0.149 | -0.215 | |
| RACEGROUP3 | 0.010 | 0.138 | 0.072 | |
| RACEGROUP4 | -0.422 | 0.159 | -2.658 | ** |
| RACEGROUP6 | -0.056 | 0.268 | -0.210 | |
| RACEGROUP7 | -0.144 | 0.142 | -1.008 | |
| Pell | 0.108 | 0.097 | 1.106 | |
| firstg | -0.183 | 0.106 | -1.736 | . |
| YRSTUDY2 | -0.084 | 0.091 | -0.928 | |
| Random Effects | Variance | Std. | | |
| case | 0.299 | 0.547 | | |
| Residual | 0.458 | 0.677 | | |
| ICC | 0.395 | | | |
| AIC | 1245.425 | | | |
| BIC | 1316.521 | | | |

 

We then added students' college conference experience (DCONF), lab research experience (DGRNOP) and the faculty mentoring construct (Mentoring) as covariates into the model. The results were similar to the previous model, and in addition, showed that the faculty mentoring was positively correlated with the scientific self-efficacy. With this model, we performed sensitivity analysis, and our preliminary results indicate that even under the condition that there exist any strong confounders, the intervention effect still holds to be strong and positive (Figure 3).

Table 5: Time Series Model with Post-test Covariates

| | Estimate | Std. Error | t statistics | sig.level |
|---|---|---|---|---|
| (Intercept) | 0.022 | 0.637 | 0.034 | |
| scholar | 0.504 | 0.228 | 2.206 | * |
| time | -0.369 | 0.085 | -4.351 | *** |
| Major1 | 0.362 | 0.230 | 1.577 | |
| Major2 | 0.559 | 0.195 | 2.864 | ** |
| HSGPA | 0.109 | 0.041 | 2.692 | ** |
| SEX | -0.228 | 0.096 | -2.369 | * |
| RACEGROUP2 | -0.012 | 0.148 | -0.081 | |
| RACEGROUP3 | 0.010 | 0.137 | 0.071 | |
| RACEGROUP4 | -0.394 | 0.157 | -2.508 | * |
| RACEGROUP6 | -0.044 | 0.265 | -0.167 | |
| RACEGROUP7 | -0.132 | 0.141 | -0.939 | |
| Pell | 0.108 | 0.096 | 1.118 | |
| firstg | -0.188 | 0.104 | -1.804 | . |
| YRSTUDY2 | -0.075 | 0.090 | -0.840 | |
| DCONF | 0.023 | 0.221 | 0.102 | |
| DGRNOP | -0.252 | 0.183 | -1.377 | |
| Mentoring | 0.176 | 0.070 | 2.509 | * |
| | | | | |
| Random Effects | Variance | Std. | | |
| case | 0.280 | 0.529 | | |
| Residual | 0.461 | 0.679 | | |
| ICC | 0.378 | | | |
| AIC | 1243.76 | | | |
| BIC | 1327.402 | | | |

 

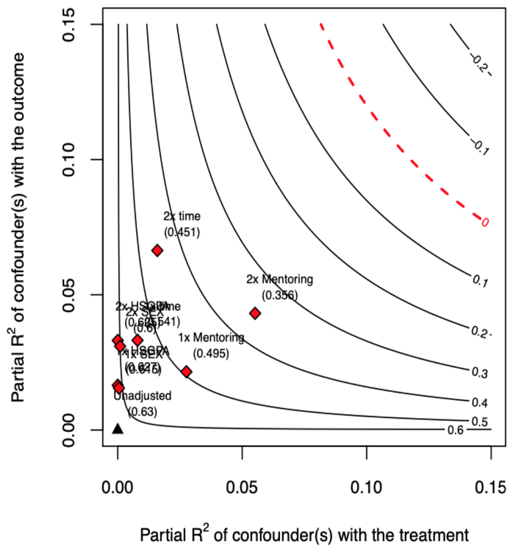


Figure 3

**Conclusions and Discussion.** Overall, preliminary findings suggest that the participation in the BUILD scholar intervention during students' first semester in college has a positive impact on science self-efficacy. In addition to the observed intervention effect, we also noticed that students' science self-efficacy tends to reduce as they progress in college. Some pre- college characteristics matter, as we found that students with higher high school GPAs are more likely to have higher science self-efficacy at the end of their first year of college. Women students, students who identified as Latino/a, and first-generation college students tend to be less likely to have high scientific self-efficacy.

With respect to college aspirations and first-year experiences, students who are biomedical natural science majors, and students who receive more positive faculty mentoring tend to have higher scientific self-efficacy. These findings suggested that the BUILD scholar program is likely to be effective, especially to students in biomedical science majors, in enhancing one's beliefs in their ability to engage in various scientific behaviors. Additionally, findings suggest that underrepresented groups, such as women, Latina/x/o students, and first-generation college students, may still need additional support to foster their science self-efficacy.

Since this study examined a cohort of BUILD students and non-BUILD students with matching baseline characteristics to determine the impacts of science self-efficacy during the first year of college, future research can focus on understanding how science self-efficacy changes over a longer period of time. The analytical models are designed to fit future time series analyses to track long-term effects of the BUILD programs. Future research can also examine the role of faculty, graduate students, and peer mentors in fostering science self-efficacy for undergraduate protégés.

**Acknowledgments.** Work reported in this publication was supported by the National Institutes of Health Common Fund and Office of Scientific Workforce Diversity (USA) by U54GM119024 and U54GM119024-03-S1 administered by the National Institute of General Medical Sciences (NIGMS).

## References

Amelink, C. T., Artis, S., & King Liu, T. J. (2015). Examining the self-efficacy of community college STEM majors: Factors related to four-year degree attainment. Community College Journal of Research and Practice, 39(12), 1111-1124.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. Psychological Review, 84(2), 191-215.

Bandura, A. (1991). Social cognitive theory of self-regulation. Organizational Behavior and Human Decision Processes, 50(2), 248-287.

Bandura, A., & Adams, N. E. (1977). Analysis of self-efficacy theory of behavioral change. Cognitive Therapy and Research, 1(4), 287-310.

Byars, A. M., & Hackett, G. (1998). Applications of social cognitive theory to the career development of women of color. Applied and Preventive Psychology, 7(4), 255- 267.

Carpi, A., Ronan, D. M., Falconer, H. M., & Lents, N. H. (2017). Cultivating minority scientists: Undergraduate research increases self-efficacy and career ambitions for underrepresented students in STEM. Journal of Research in Science Teaching, 54(2), 169-194.

Chemers, M. M., Hu, L. T., & Garcia, B. F. (2001). Academic self-efficacy and first year college student performance and adjustment. Journal of Educational Psychology, 93(1), 55-64.

Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. Educational Research Review, 17, 63-84.

Lent, R. W., Brown, S. D., & Hackett, G. (2000). Contextual supports and barriers to career choice: A social cognitive analysis. Journal of Counseling Psychology, 47(1), 36–49.

MacPhee, D., Farro, S., & Canetto, S. S. (2013). Academic self-efficacy and performance of underrepresented STEM majors: Gender, ethnic, and social class patterns. Analyses of Social Issues and Public Policy, 13(1), 347-369.

 