**Critique of**

“Examining the Effects of Student Involvement

on African American College Student Development”

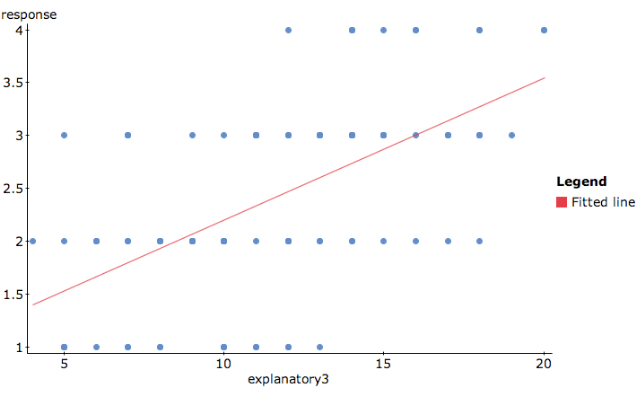
These are graded on a 4-point scale. As specified in the assignment, I expected you to address (at least) the questions below. Grades were based on the correctness of your critique (for example, if you said that he *did* check his conditions, you did not get points for answering that question correctly). I did not expect you to pick out all the details or mention every issue that I do below, but I did expect you to address every question in some way, and to come to the same general conclusions.

(For this particular article, the biggest problem is the use of ordinal data. If you did not mention that issue, you could not earn a 4.)

* Is a linear regression model appropriate in this context?

No. The use of Likert scales (ordinal data) throws the entirety of the analysis into question. Likert scales are, by definition, categories that have had numbers assigned to them. This does not magically make them numerical variables. Think about it like this: if the variables are (for example) “length” and “width”, a length of 2 meters is twice a length of 1 meter, and a length of 4 meters is twice 2 meters. But when you use a Likert scale, is “occasionally” (2) twice as often as “never” (1)? Is “very much” (4) twice as much as “some” (2)? The ratios don’t make sense because these aren’t really numbers, they’re *categories*. Thus, linear regression isn’t appropriate here.

For further explanation: Below is a scatterplot of data with a 1-4 Likert-scale response variable and explanatory variable that ranges from 4 to 20 (integers only), which is how Dr. Flowers’ data is set up. The regression line is added to this plot. This regression model has an R2 of 34%, which is significantly higher than all the R2’s in this article. Does it really look like a linear regression model is appropriate here?



* Does the author discuss any diagnostic plots or mention whether the assumptions of the linear regression model are met?

No. The author did not check any of the assumptions/conditions for the simple linear regression model. At least, he didn’t make any mention about the conditions, or any mention of him checking them; thus, we have no information that the conditions are met.

* Is the sample representative of the larger population about which he makes conclusions?

It’s questionable that his data could be generalized to all African-American college students. Nearly 40% of his sample are freshmen, and since he’s asking about college experiences and gains, freshmen just aren’t going to have as much to say as upperclassmen. Another problem is the potential for self-selection bias: the students who feel excited or compelled to answer this questionnaire may be “joiner” types, and thus more likely to be involved in their institutions.

There were nearly twice as many women as men, which seems like it can’t be representative. However, the latest Dept of Education figures show that black women account for 63.6% of all African American enrollments[[1]](#footnote-1). (Thanks to Zach for finding this figure and including the source!) So the gender distribution is actually okay (although if you mentioned it as a potential issue, that’s fine; I found it suspect too until I saw Zach’s figure).

One more note: Several of you mentioned that because the sample wasn’t random, it couldn’t possibly be representative. That’s not actually true. Representativeness is what we want. The reason there’s so much focus on “random sampling” is because a random sample is most likely to be representative. But there can be (and are) non-random samples that are representative, and that’s okay because representation is what matters at the end of the day.

* What inference methods did he perform?

He performs 410 separate simple linear regressions, between each of his explanatory and each of his response variables. For each regression, he performs a t-test on the slope (we assume H0: slope = 0, since that’s standard), and flags it as significant if the p-value is less than 0.001.

He also reports R2 values, but I have no idea what those R2‘s are from. He puts them at the bottom of each table as if they’re for a multiple linear regression with all the variables in that column, but that can’t be right because he specifically states (p. 638) that he’s doing individual simple (not multiple) linear regressions, which means there should be 410 R2 values. He also labels the R2‘s as significant (with p-values), which is nonsensical: there is *no statistical test* for R2.

* Are his conclusions valid for the tests that he performed?

No. The fact is, *if the conditions of the linear model are not met, the linear regression model is not appropriate and any inference/conclusions he makes are meaningless.* Since we don’t know if the conditions are met, we don’t know if what he’s saying is meaningless or not. (I would argue that in this case, it’s meaningless regardless of conditions because of the use of ordinal data.)

* Other problems you may have noticed:

His interpretation of Table 1 is incorrect. He states that “...African American students reported the greatest amount of gains in the following areas: vocational preparation, thinking and writing skills, and understanding science and technology” (p.638). But from my reading of Table 1, the largest gains were actually in personal and social development, thinking and writing skills, and vocational preparation. More importantly: all categories have average gains between 2.11 (US&T) and 2.81 (P&SD) per item, so are the differences really significant?

In his interpretations, he assumes that correlation = causation. He states several times that “X *leads to* Y” or “A *results in* B” (emphasis mine). Correlation is not causation! (And we don’t know even know how correlated these things are, since he doesn’t provide any correlation coefficients. Not that it matters, because correlation is not a useful measure when dealing with ordinal data anyway.)

There are also claims made that seem to contradict themselves in the paper. In the Summary of Results section, Flowers claims, “For most of the student involvement experiences analyzed in the study, there was only a trivial and nonsignificant direct effect on student development for African American students.” (648) Presumably this refers to the fact that the R2 values were low (around 0.1 or 0.2). However, directly following this, Flowers claims in the Discussion, “Findings of this study suggest that student involvement experiences directly impact student development gains in [list of types of gains] for African American students” (648). This seems like a contradiction after Flowers stated that the direct effect was “trivial and nonsignificant.” (Credit to Margaret for pointing this out.)

1. "Black Women Students Far Outnumber Black Men at the Nation's Highest-Ranked Universities." The Journal of Blacks in Higher Education. Accessed September 13, 2018. http://www.jbhe.com/news\_views/51\_gendergap\_universities.html. [↑](#footnote-ref-1)