5/18/2022 Meeting

Sex Bias paper: discuss edits

Next Step paper

1. *Figure 1: Pairwise Correlations* numbers didn’t match 🡪 from different cycles

* Generated figures for 2019-20, 2020-21, and combined cycles
* Can create a Figure folder in the shared drive under ERAS Linguistics/Manuscripts & upload the 3 figures there

1. R and p values in correlation plots

* P-values are used in hypothesis tests to determine whether we reject or fail to reject the null hypothesis.
* For Pearson's correlation coefficient, we can perform statistical test to determine whether the correlation coefficient is significant.
* The null hypothesis is that the true correlation coefficient *ρ* is equal to 0, based on the value of the sample correlation coefficient *r*.
  + H0: ρ = 0 vs H1: ρ ≠ 0
* p-value < 0.05 🡪 reject null & can conclude that the correlation coefficient is different from zero and that a linear relationship exists.

**Formula**

The p-value for Pearson's correlation coefficient uses the t-distribution.



The p-value is 2 × P(T > t) where T follows a t distribution with *n* – 2 degrees of freedom.

Do we want to include R and p values in each panel?

Choose a Step 1 axis limit

Step 1 (min, max): (198, 270)

Step 1 axis limit (-/+ 5 from min/max): (193, 275) Step 1 axis limit (-/+ 3 from min/max): (195, 273)

Chart, scatter chart

Description automatically generated Chart

Description automatically generated

Do we want unlabeled tick marks in Panel A? First figure - without, second figure - with

Chart, scatter chart

Description automatically generated

Same x axis in all 4 panels

Chart, scatter chart

Description automatically generated

**4- vs 3-panel figure**: 4-panel figure looks the best (none of the 3-panel arrangement looks reasonable). If want emphasis on Step1 vs Step2, we can present Panel A on its own in addition to the 4x4 figure.

Chart, scatter chart

Description automatically generated

Diagram

Description automatically generated

A picture containing timeline

Description automatically generated

Statistically sig - Opposite of our original hypothesis: female gets more positive letters

But no practical meaning

FINAL: same x-axis in all 4 panels + R and p values

Chart, scatter chart

Description automatically generated

The correlation you listed (r=0.0557) tells you that you are looking at two variables that appear to have little to no collinearity. I fully agree with Ariel that for such low correlation (and numerically speaking even lower coefficient of determination) statistical significance is bound to be low.

However, statistical significance will be more dependent on the sample size than on the degree of correlation/determination. Consequently, for very large sample sizes with almost no collinearity, you may see highly statistically significant results, and vice versa (e.g., you may see very high correlation/determination without statistical significance in a small sample).

The key questions here are: (a) is correlation or significance more important? ; (b) what is the expected vs observed degree of correlation? ; (c) what is the significance of such relationship(s) in practical application of research?

That last question is perhaps the most relevant... two further examples:

1) If you happen to look at variables that may have high degree of correlation but the resultant change in process or practice makes little to no practical significance, then you should be asking the "who cares?" question

2) If you happen to look at variables with small degree of correlation, but even the slightest change in the associated process or practice will result in dramatic practical consequences, then the approach should be "can't ignore it"...

The main point here -- correlate your research results with their impact on real life processes and practices. If your research leads to any change that results in more than minimal process or practice alteration, then you should never ignore it. And please remember... change can be both positive or negative, so going from one level of correlation to another may result in either desirable or undesirable change, regardless of significance. You will not know the associated consequences until you either accurately model or actually implement said change.

Chart, scatter chart

Description automatically generated

The number of peer reviewed abstracts and the number of poster presentations have R values of

The correlation coefficient between PRA and Step 1 score is -0.00072

Of all the

The near-zero correlation coefficients that PRA and PP have with Step 1 score— R= -0.00072 and R = -0.03, respectively—indicate a lack of linear relationship.

Graphical user interface, text, application

Description automatically generated

Text

Description automatically generated

R() = Text

Description automatically generated Text

Description automatically generated

Nine hundred and twenty-eight (n=928) applications were examined for their Step 1 scores (M = 245.2, SD = 12.68, NA = 7) and their correlation with nine other metrics in the applications. The relationship between Step 1 score and four other metrics—including Step 2 scores (M =253, SD = 12.37, NA = 130), number of peer reviewed articles (M = 3.17, SD = 3.82), number of oral presentations (M = 2.54, SD = 2.85), and number of posters (M = 4.542, SD = 4.4)—were evaluated in detail. See Figure XXX for their scatterplots.

Pearson’s r data analyses revealed Step 1 and Step 2 scores have moderate, positive, and statistically significant linear relationship (r(796) = .67, p < .001), suggesting that that students tend to perform consistently across the two exams. On the other hand, Step 1 scores had little to no collinearity with the number of peer reviewed abstracts (r(919) = -.00072, p = .982), the number of oral presentations (r(919) = -.072, p = .029), or the number of poster presentations (r(919) = -.03, p=.361). The small Pearson’s r in all three cases means that one cannot reliably predict a student’s output/standing in any of the three metrics given Step 1 scores. In fact, there was practically no collinearity between Step 1 score and all the other metrics we analyzed except for Step 2 score. While Step 2 score can be used to predict a student’s corresponding Step 1 score to a moderate extent, the correlation is not high enough—or, the relationship is not strong enough—for Step 2 score to serve as a replacement for Step 1 score.

what is the significance of such relationship(s) in practical application of research?

Text

Description automatically generatedText

Description automatically generatedText

Description automatically generatedText

Description automatically generated

A screenshot of a computer

Description automatically generated with low confidence Text

Description automatically generated