CONFIDENTIAL

FSU Case # RM-CQ99

Date: 11/03/2021

Florida State University
Inquiry Report Concerning
Allegations of Research Misconduct against
Eric Stewart

1. Name and position of the Respondents

Eric Stewart, Professor, Criminology

2. Description of the allegations of research misconduct

"The main finding is that across the 10 papers the reported SDs for many, but not all, binary variables are incorrect." (Excerpt from complaint document named ~ Eric Stewart Paper Standard Deviations)

3. The external support pertinent to the allegation

Federal Award: NIMH, CDCP, NIDA, NIAAA

4. The names and titles of the committee members and experts who conducted the inquiry

- Kathryn Tillman, Professor, Sociology
- Fred Huffer, Professor, Statistics
- Debajyoti Sinha, Professor, Statistics

5. Summary of the inquiry process used

The committee members reviewed the materials provided in the complaint. In lieu of an interview, the committee addressed three sets of queries to the respondent and held three committee meetings (via Zoom) to review the responses received, discuss the evidence, and form a concensus recommendation.

6. List of the research records reviewed

(1) Pdf files for eleven articles and three additional references all of which included the respondent as a co-author, (2) a spreadsheet containing a summary statement and a list of the means, SD's, and their discrepancies for binary variables used in the eleven articles, and (3) partial re-creations of the data used in the Berg et al. 2016 and Warren et al. 2020 articles. Items (1) and (2) were included with the complaint, and (3) was supplied by the respondent.

7. Summaries of Respondent Interview(s)

In its first two sets of queries the committee learned that the respondent no longer had the code (due to a hard disk crash) or access to the data (due to data use restrictions) that was used in the various articles listed in the complaint. (We did receive partial re-creations of the data used in two of the articles.) In answering the third set of queries, the respondent (1) stated that he had received no correspondence from the journals relating to mistakes or anomalies in the articles listed in the complaint, and (2) described the typical approach he uses in his research for dealing with missing data.

8. Committee recommendation and the basis for the recommendation

For binary variables, that is, variables which assume only the values 0 and 1, there is a mathematical relationship between and the mean and standard deviation (SD). The complaint notes that, in 11 articles on which the respondent is a co-author, many of the reported means and SD's for binary variables fail to satisfy this relationship. The complaint lists these cases and gives the discrepancy between the reported SD and the value expected by the relationship. There are three mistakes in this listing: the variables Heterogeneity of College Goals in Article 2 (Berg et al. 2013) and Violent Victimization (T1) amd Violent Victimization (T2) in Article 9 (Stewart, Schreck, and Simons 2006) are not binary variables and thus should not be included in the listing. There are also two cases in the list which are very likely simple transcription errors (misplaced decimal places), and a few cases where the discrepancy is small enough to be due to round-off errors. But even when all these cases are excluded, there are still sufficiently many genuinely anomalous SD's to indicate a definite problem. Our committee believes it is unlikely that this many anomalous SDs could be entirely due to rounding, transcription, and copy-editing errors. Moreover, the anomalous SD's exhibit a pattern; in those cases where the discrepancy is too large to be due to round-off errors, nearly all the anomalous SDs are too small. Transcription errors are unlikely to exhibit this pattern.

So, we agree with the statement in the complaint that there are errors in many of the reported SD's for binary variables in the listed articles. However, these SD's are not important in themselves; they have no direct relationship with the main conclusions in these articles, and there is no obvious motivation for the respondent to alter or fabricate these particular numbers.

Moreover, these anomalous SD's can be explained by means which do not involve any research misconduct. In particular, the treatment of missing values, especially if this is done improperly, can result in anomalous SD's for binary variables. For example, if a binary variable has missing values, and these are replaced by the mean of the non-missing values (mean imputation), then the computed SD will be too small. This is exactly what is observed in the vast majority of the anomalous SD's mentioned in the complaint: the SD's are too small. More sophisticated regression imputation methods can also produce SD's that are too small. The respondent stated that he typically uses "multiple imputation" to deal with missing values, even in the computation of the summary statistics (the mean and SD). Multiple imputation methods are less likely to underestimate the variance, but if applied incorrectly (say, by using linear regression instead of logistic regression for the imputation of binary variables), they can still result in means and SD's which fail to satisfy the relationship expected for binary variables.

Paper 1 (Berg et al 2016) used data from the Family and Community Health Study (FACHS) and states on page 529 that "Approximately 3 percent of the sample had missing data on items used in our analyses" and further states in a footnote that multiple imputations by chained equations in Stata (a statistical software package) was used to deal with these missing values. None of the other papers listed in the complaint mentioned anything about missing values, but since nine of these articles used data from the same source (FACHS), it is likely that missing values existed in the data for many (or all) of these articles and that the respondent used some form of imputation for dealing with them. Therefore, problems with the imputation of missing values could potentially explain all of the anomalous SD's.

Our committee had wished to examine the code and data used in some of the articles to determine the precise cause of the anomalous SD's, but learned that the code had been lost in a hard disk crash and that the data was no longer available (due to data use restrictions). There is the potential for surviving code in the respondent's sequestered computer equipent or in the crashed hard disk (if it still exists), but based on the evidence the committee has seen so far,

there is little reason to believe the anomalous SD's are due to research misconduct as opposed to honest error. For this reason we believe that a full investigation is not warranted.

9. Respondent comments on the draft report

☐ **are attached, or** ☒ **the respondent chose not to provide any comments.**

10. Whether any other actions should be taken if an investigation is not recommended

Report submitted by (name and signatures of all Committee members)

_____

_____

_____

FLORIDA STATE UNIVERSITY | OFFICE *of the* VICE PRESIDENT *for* RESEARCH

TO: Laurel Fulkerson, Deciding Official
Interim Vice President for Research

FROM: Diana Key, Research Integrity Officer (RIO)
Director, Research Compliance Programs

Attached is the final report from the Inquiry Committee. Please review this document and, as Deciding Official, determine next steps in this case. Please mark your response below as appropriate.

☐ I accept the findings and recommendations of the Inquiry Committee, and direct the RIO to proceed in accordance with FSU policy and procedures.

☐ I do not accept the findings and recommendations of the Inquiry Committee. I direct the RIO to reconvene the Inquiry Committee for further fact-finding and analysis as follows:

☐ I override the findings and recommendations of the Inquiry Committee with the following determination/decision:

Deciding Official Signature/Date:_____