Over the last decade, it has become clear that many published research articles contain errors in the reported statistical results (Bakker & Wicherts, 2011; Berle & Starcevic, 2007; Caperos & Pardo, 2013; Garcia-Berthou & Alcaraz, 2004; Nuijten, Hartgerink, van Assen, Epskamp, & Wicherts, 2015; Veldkamp, Nuijten, Dominguez-Alvarez, van Assen, & Wicherts, 2014; Wicherts, Bakker, & Molenaar, 2011). Veldkamp et al. (2014) investigated whether collaboration on statistical analyses, as proposed by Wicherts (2011), was associated with reduced error rates in psychology articles. While they confirmed the high reporting error rates in psychology and established that collaboration on statistical analysis is uncommon among psychologists, they found no association between collaboration and reporting error rates. However, their data had several limitations that may have limited the study’s capability to detect this potential association. First, estimation of collaboration on statistical analysis was based on a questionnaire asking authors to recall who had been involved in the statistical analyses reported in their paper, which may have yielded inaccurate responses. Second, e-mails from the respondents to the authors in response to the survey showed that the questionnaire was perceived as sensitive, possibly leading to response bias and expectancy effects. Third, collaboration was defined as ‘more than one author involved in the statistical analyses’, prohibiting analysis of the relationship between reporting errors and *number* of authors involved in the analysis. Fourth, the study suffered from lack of power due to a low overall response rate and an overestimated expected effect size.

For these reasons, we propose an alternative design to test the hypothesis that collaboration on statistical analysis can help reduce statistical reporting errors: by relating statistical reporting errors to the number of authors responsible for the data analysis as reported in the ‘author contributions’ section of all psychology research articles published in the multidisciplinary journal PLoS One to date (currently 12.098). This does not only allow us to more objectively determine how many authors were involved in the statistical analyses, but also to examine in a much much larger sample whether there is an association between the number of authors involved in the analyses and error rates. While our software to detect statistical reporting can currently only be applied to articles employing the reporting style of the American Psychological Association (APA) style (American Psychological Association, 2010) we will also extract the author contribution information from all other research articles published in PLoS One to date (currently 149.460). This will enable us to compare major research areas in terms of mean number of authors responsible for the statistical analyses relative to the number of authors on the article, how often the first author of the paper is listed among those responsible for the data analysis, and how many of the authors who contributed to writing the paper are listed among those responsible for the data analysis.

**1A. Main hypothesis**

1. Collaboration on statistical analysis is associated with reduced error probabilities in the reporting of statistical results in published psychology research articles.

**1B. Main statistical hypotheses to test confirmatively:**

1. The probability that a paper contains at least one p-value that is incongruent with its test statistic and degrees of freedom is lower in papers in which more than one person was involved in the statistical analyses than in papers in which only one person was involved in the statistical analyses.  *Simple logistic regression analysis with whether a paper contains at least one p-value that is congruent as dependent variable and (dummy-coded) number of authors involved in the analyses as predictor.*
2. The probability that a paper contains at least one p-value that is incongruent with its test statistic and degrees of freedom to the extent that it affects statistical significance is lower in papers in which more than one person was involved in the statistical analyses than in papers in which only one person was involved in the statistical analyses. *Simple logistic regression analysis with whether a paper contains at least one p-value that is incongruent to the extent that is affects statistical significance as dependent variable and (dummy-coded) number of authors involved in the analyses as predictor.*
3. The probability that a p-value is incongruent with its test statistic and degrees of freedom is lower if more than one person was involved in the statistical analyses than when only one person was involved in the statistical analyses. *Multilevel logistic regression analysis with whether a p-value is incongruent as dependent variable, (dummy-coded) number of authors involved in the analyses as fixed factor, and article as random factor.*
4. The probability that a p-value is incongruent with its test statistic and degrees of freedom to the extent that it affects statistical significance is lower if more than one person was involved in the statistical analyses than when only one person was involved in the statistical analyses. *Multilevel logistic regression analysis with whether a p-value is incongruent to the extent that it affects statistical significance as dependent variable, (dummy-coded) number of authors involved in the analyses as fixed factor, and article as random factor.*

**2A. Secondary hypotheses**

1. The first author being involved in the statistical analyses is associated with increased serious error probabilities in the reporting of statistical results in published psychology research articles.
2. Reported conflicts of interest are associated with increased serious error probabilities in the reporting of statistical results in published psychology research articles.

**2B. Secondary statistical hypotheses to test confirmatively:**

1. The probability that a paper contains at least one p-value that is incongruent with its test statistic and degrees of freedom to the extent that it affects statistical significance is higher in papers in which the first author was involved in the statistical analyses than in papers in which the first author was not involved in the statistical analyses. *Simple logistic regression analysis with whether a paper contains at least one incongruent p-value as dependent variable and whether the first author was involved in the analyses as predictor.*
2. The probability that a p-value is incongruent with its test statistic and degrees of freedom to the extent that it affects statistical significance is higher in papers in which the first author was involved in the statistical analyses than in papers in which the first author was not involved in the statistical analyses. *Multilevel logistic regression analysis with whether a p-value is incongruent as dependent variable, whether the first author was involved in the statistical analyses as fixed factor, and article as random factor.*
3. The probability that a paper contains at least one p-value that is incongruent with its test statistic and degrees of freedom to the extent that it affects statistical significance is higher in papers in which the authors declared a conflict of interest than in papers in which the authors declared no conflict of interest. *Simple logistic regression analysis with whether a paper contains at least one incongruent p-value as dependent variable and whether the declared a conflict of interest as predictor.*
4. The probability that a p-value is incongruent with its test statistic and degrees of freedom to the extent that it affects statistical significance is higher in papers in which the authors declared a conflict of interest than in papers in which the authors declared no conflict of interest. *Multilevel logistic regression analysis with whether a p-value is incongruent as dependent variable, whether the authors declared a conflict of interest as fixed factor, and article as random factor.*

**3A. Exploratory research questions**

1. Are there differences between countries in error probabilities in the reporting of statistical results in published psychology research articles?

**3B. Exploratory statistical analyses**

1. The probability that a paper contains at least one p-value that is incongruent with its test statistic and degrees of freedom can be predicted by country. *Simple logistic regression analysis with whether a paper contains at least one incongruent p-value as dependent variable and country as predictor.*
2. The probability that a paper contains at least one p-value that is incongruent with its test statistic and degrees of freedom to the extent that it affects statistical significance can be predicted by country. *Simple logistic regression analysis with whether a paper contains at least one p-value that is incongruent to the extent that is affects statistical significance as dependent variable country as predictor.*
3. The probability that a p-value is incongruent with its test statistic and degrees of freedom can be predicted by country. *Multilevel logistic regression analysis with whether a p-value is incongruent as dependent variable, country as fixed factor, and article as random factor.*
4. The probability that a p-value is incongruent with its test statistic and degrees of freedom to the extent that it affects statistical significance can be predicted by country. *Multilevel logistic regression analysis with whether a p-value is incongruent to the extent that it affects statistical significance as dependent variable, country as fixed factor, and article as random factor.*

**Data**

Data to collect

*All research articles PLoS One*

* Doi
* Authors (full names, in order of authorship)
* Author contributions, 4 sections [author\_notes: ‘conceived and designed the experiments’, ‘performed the experiments’, ‘analyzed the data’ and ‘wrote the paper’] (initials).
* Subject area level 1, all categories listed
* Subject are level 2, all categories listed *(psychology should be level 2 as its level 1 is ‘social sciences’)*

*For all psychology articles, also:*

* Results section
* Subject area level 3 (i.e. within psychology), all categories listed

Data to extract from papers

* List of author names [order must be useable to determine author position] *(all research papers)*
* First author name
* Country of first author (i.e. country of (1st) affiliation of first author) *(all research papers)*
* List of abbreviated author names (initials first name and surname) [create based on author names, test if created initials appear in any of the extracted author contributions, if not, check. Again order must be useable to determine author position] *(all research papers)*
* List of abbreviated author names of those who ‘Conceived and designed the experiments’ [in author contributions 1] *(all research papers)*
* List of abbreviated author names of those who ‘Performed the experiments’ [in author contributions 2] *(all research papers)*
* List of abbreviated author names of those who ‘Analyzed the data’ [in author contributions 3] (CHECK!! Is this always nr 3??) *(all research papers)*
* List of abbreviated author names of those who ‘Wrote the paper’ [in author contributions 4] *(all research papers)*
* Subject areas [first level subject areas, string with all subject areas] *(all research papers)*
* Subject areas [second level subject areas, string with all subject areas] *(all research papers)*
* Subject areas [third level subject areas, string with all subject areas] *(psychology papers only)*
* Results section *(psychology articles only)*
  + Extract from results section:
    - All test-statistics
    - All *p*-values
    - Whether *p*-values are consistent or not
    - Whether inconsistent *p*-value is a gross error or not

Variables to compute and/or use for analyses:

* Count number of authors on paper [number] *(all research papers)*
* Count number of authors who conceived and designed the experiments [number] *(all research paper)*
* Count number of authors who performed the experiments [number] *(all research papers)*
* Count number of authors who analyzed the data [number] *(all research papers)*
* Count number of authors who wrote the paper [number] *(all research papers)*
* Count proportion of authors who conceived and designed the experiments (relative to number of authors on paper) [%?] *(all research paper)*
* Count proportion of authors who performed the experiments (relative to number of authors on paper) [%?] *(all research papers)*
* Count proportion of authors who analyzed the data (relative to number of authors on paper) [%?] *(all research papers)*
* Count proportion of authors who wrote the paper (relative to number of authors on paper) [%?] *(all research papers)*
* Indicate whether the first author on the paper is also the first contributor in the list of authors who wrote the paper [0/1] *(all research papers)*
* Indicate whether the first author on the paper is the first contributor in the list of authors who analyzed the data [0/1] *(all research papers)*
* Indicate whether the first author on the paper is listed at all among the authors who analyzed the data[0/1] *(all research papers)*
* Names subject area level 1 [list/factor?] *(all research papers)*
* Names subject area level 2 [list/factor?] *(all research papers)*
* Names subject area level 3 [list/factor?] *(psychology papers only)*
* Name of country of first affiliation of first author
* Number of *p*-values per paper [number] *(psychology papers only)*
* Number of inconsistent *p*-values per paper [number] *(psychology papers only)*
* Number of gross errors per paper [number] *(psychology papers only)*

American Psychological Association. (2010). *Publication Manual of the American Psychological Association. Sixth Edition*. Washington, DC: American Psychological Association.

Bakker, M., & Wicherts, J. M. (2011). The (mis)reporting of statistical results in psychology journals. *Behavior Research Methods, 43*, 666-678. doi: 10.3758/s13428-011-0089-5

Berle, D., & Starcevic, V. (2007). Inconsistencies between reported test statistics and p-values in two psychiatry journals. *International Journal of Methods in Psychiatric Research, 16*(4), 202-207. doi: 10.1002/mpr.225

Caperos, J. M., & Pardo, A. (2013). Consistency errors in p-values reported in Spanish psychology journals. *Psicothema, 25*(3), 408-414. doi: 10.7334/psicothema2012.207

Garcia-Berthou, E., & Alcaraz, C. (2004). Incongruence between test statistics and P values in medical papers. *Bmc Medical Research Methodology, 4*, 13. doi: 10.1186/1471-2288-4-13

Nuijten, M. B., Hartgerink, C. H., van Assen, M. A., Epskamp, S., & Wicherts, J. M. (2015). The prevalence of statistical reporting errors in psychology (1985–2013). *Behavior Research Methods*, 1-22.

Veldkamp, C. L., Nuijten, M. B., Dominguez-Alvarez, L., van Assen, M. A., & Wicherts, J. M. (2014). Statistical reporting errors and collaboration on statistical analyses in psychological science. *PLoS One, 9*(12), e114876.

Wicherts, J. M. (2011). Psychology must learn a lesson from fraud case. *Nature, 480*, 7. doi: 10.1038/480007a

Wicherts, J. M., Bakker, M., & Molenaar, D. (2011). Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results. *PLoS One, 6*(11), e26828. doi: 10.1371/journal.pone.0026828