Supplement 4. Disproportionate stratified random cluster sample of the population of research articles and reports in *American Antiquity*, *Journal of Archaeological Science*, *Journal of Field Archaeology*, and *Journal of Archaeological Method and Theory* from January 2000 to December 2019, summarizing the counts and proportions of articles that use “sample” or sampling in various ways, along with evaluation of the stratification’s effectiveness.

The sampling involved randomly selecting 20% of the published issues by *American Antiquity*, *Journal of Archaeological Science* (*JAS*), and *Journal of Archaeological Method and Theory* (*JAMT*) in this period, but only 5% of those in *Journal of Archaeological Science* (*JAS*), which has a much greater frequency of publication than the others. These four journals were the strata. The number of clusters (*n*) in each stratum was 16 (*American Antiquity* and *JAMT*), 17 (*JFA*), and 12 (*JAS*), while the total number of articles sampled in each stratum was 144 (*American Antiquity*), 167 (*JAS*), 103 (*JFA*), and 108 (*JAMT*), with a total of 522 articles. This is a cluster sample, consisting of every research paper that appears in each sample issue (by contrast, a multistage sample would take a subsample from each sample issue), so each stratum’s variance is calculated from the squared differences of each issue’s proportion from the overall cluster proportion. As the sampling fractions are unequal (disproportionate stratified sampling), the summary statistics to estimate characteristics of the whole population require weighting factors that are proportional to the size of each stratum’s subpopulation.

Table S4.1. Each row shows, by stratum and population, the number of articles in the sample and the estimated cluster percentages in the population that discuss samples or sampling in the way indicated (“sampling concerns”). The error ranges (±) are cluster Standard Errors, and suspected outliers are bolded. Standard Errors vary on topics with the same number of articles because of variation in the number of articles per issue.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***Am Antiq*** | | ***Jn Arch Sci*** | | ***Jn Field Arch*** | | ***JAMT*** | | **All** | |
| **Sampling Concern** | **No** | **%** | **No** | **%** | **No** | **%** | **No** | **%** | **No** | **%** |
| 1. Unspecified sampling method | 45 | **31.3** ± 5.3 | 44 | 26.3 ± 4.2 | 22 | 21.4 ± 4.4 | 14 | **12.5** ± 2.8 | 125 | 24 ± 1.1 |
| 2. Convenience sampling | 2 | 1.4 ± .8 | 1 | .6 ± .33 | 2 | 1.9 ± 1.5 | 0 | 0 | 5 | .9 ± .2 |
| 3. Survey “sampling” w/o sampling theory | 13 | 9.0 ± 2.5 | 29 | 17.4 ± 3.5 | 21 | 20.4 ± 3.9 | 6 | **5.4** ± 1.6 | 69 | 14.5 ± 8.6 |
| 4. “Full-coverage” survey | 0 | 0 | 0 | 0 | 2 | 1.9 ± 1.8 | 0 | 0 | 2 | .3 ± .2 |
| 5. “Sample” is whole population | 14 | 9.7 ± 2.5 | 17 | 10.2 ± 2.7 | 9 | 8.7 ± 3.4 | 3 | **2.7** ± 1.1 | 43 | 8.6 ± .7 |
| 6. Claimed “representative” sample | 3 | 2.1 ± 1.0 | 3 | 1.8 ± 1.3 | 7 | **6.8** ± 2.3 | 3 | 2.7 ± 1.2 | 16 | 2.8 ± .4 |
| 7. Discusses sample size only | 3 | 2.1 ± 1.0 | 1 | .6 ± .55 | 0 | 0 | 0 | 0 | 4 | .6 ± .3 |
| 8. Experiments w/o randomization | 1 | **.7** ± .7 | 18 | 10.8 ± 2.7 | 7 | 6.8 ± 2.5 | 13 | 11.6 ± 3.2 | 39 | 8.6 ± .7 |
| 9. “Sampling” for “collecting” | 4 | 2.8 ± 1.5 | 1 | .6 ± .55 | 3 | 2.9 ± 1.3 | 0 | 0 | 8 | 1.3 ± .25 |
| 10. “Systematic” for “methodical” | 1 | .7 ± .7 | 0 | 0 | 0 | 0 | 2 | 1.8 ± .9 | 3 | .4 ± .1 |
| 11. “Carbon samples” | 41 | **28.5** ± 3.8 | 17 | 10.2 ± 2.3 | 26 | 25.2 ± 3.5 | 8 | **7.1** ± 2.9 | 92 | 15.3 ± .7 |
| 12. Bone specimens | 7 | 4.9 ± 1.6 | 28 | **16.8** ± 3.2 | 4 | 3.9 ± 2.5 | 6 | 5.4 ± 3.5 | 45 | 10.7 ± .8 |
| 13. “Sample” = Lithic specimen | 5 | 3.5 ± 1.2 | 5 | 3 ± 1.5 | 5 | 4.9 ± 4.1 | 1 | **.9** ± .7 | 16 | 3.0 ± .5 |
| 14. “Sample” = Pottery specimen | 11 | 7.6 ± 2.6 | 7 | 4.2 ± 2.4 | 8 | 7.8 ± 3.2 | 4 | 3.6 ± 4.7 | 30 | 5.3 ± .7 |
| 15. Flotation “samples” | 11 | 7.6 ± 1.6 | 6 | 3.6 ± 1.0 | 8 | 7.8 ± 3.5 | 1 | **.9** ± .5 | 26 | 4.5 ± .4 |
| 16. “Soil samples” | 16 | 11.1 ± 2.8 | 29 | **17.4** ± 2.8 | 29 | 28.2 ± 4.5 | 5 | **4.5** ± 2.1 | 79 | 16 ± .8 |
| 17. Other specimens | 12 | 8.3 ± 2.1 | 49 | **29.3** ± 4.3 | 23 | 22.3 ± 5.4 | 5 | 4.5 ± 3.4 | 89 | 20.5 ± 1.1 |
| 18. Explicitly neglects cluster formulae | 2 | 1.4 ± .9 | 0 | 0 | 1 | 1.0 ± 1.0 | 0 | 0 | 3 | .4 ± .1 |
| Total | 191 |  | 255 |  | 177 |  | 71 |  | 694 |  |

The next four tables provide detailed information on the content of each stratum.

Table S4.2. Random cluster sample without replacement from the *American Antiquity* stratum. Key to column headings is in the left column of table S4.1.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Year | Vol | Iss | Articles | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| 2001 | 66 | 4 | 11 | 6 | 1 | 0 | 0 | 0 | 1 | 3 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 3 | 0 | 0 |
| 2002 | 67 | 3 | 7 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| 2003 | 68 | 2 | 8 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2004 | 69 | 1 | 8 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 3 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 2004 | 69 | 2 | 10 | 3 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 |
| 2005 | 70 | 4 | 8 | 5 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 2 | 0 | 0 |
| 2008 | 73 | 1 | 7 | 5 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2008 | 73 | 3 | 8 | 3 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 3 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| 2010 | 75 | 1 | 9 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 3 | 1 | 1 | 1 | 0 |
| 2011 | 76 | 4 | 8 | 2 | 0 | 0 | 0 | 3 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 |
| 2013 | 78 | 1 | 10 | 2 | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| 2014 | 79 | 4 | 12 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 6 | 2 | 1 | 0 | 0 | 0 | 1 | 0 |
| 2015 | 80 | 2 | 11 | 1 | 1 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 3 | 1 | 1 | 3 | 0 |
| 2017 | 82 | 4 | 9 | 4 | 0 | 0 | 0 | 2 | 1 | 0 | 0 | 1 | 0 | 2 | 0 | 1 | 1 | 1 | 3 | 2 | 1 |
| 2018 | 83 | 2 | 9 | 1 | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 4 | 1 | 0 | 1 | 2 | 2 | 1 | 0 |
| 2019 | 84 | 3 | 9 | 2 | 0 | 1 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 1 | 0 | 1 | 1 | 2 | 0 |
| Total | | 16 | 144 | 45 | 2 | 13 | 0 | 14 | 3 | 3 | 1 | 4 | 1 | 41 | 7 | 5 | 11 | 11 | 16 | 12 | 2 |

Table S4.3. Random cluster sample without replacement from the *Journal of Archaeological Science* stratum. Key to column headings is in the left column of table S4.1.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Year** | **Vol** | **Iss** | **Articles** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** |
| 2000 | 27 | 12 | 10 | 5 | 0 | 1 | 0 | 0 | 0 | 0 | 3 | 1 | 0 | 0 | 2 | 1 | 0 | 0 | 3 | 2 | 0 |
| 2001 | 28 | 9 | 10 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 4 | 0 |
| 2003 | 30 | 11 | 13 | 4 | 0 | 4 | 0 | 2 | 1 | 1 | 2 | 0 | 0 | 4 | 6 | 1 | 0 | 1 | 2 | 5 | 0 |
| 2003 | 30 | 12 | 15 | 3 | 0 | 2 | 0 | 2 | 2 | 0 | 3 | 0 | 0 | 2 | 1 | 0 | 0 | 0 | 2 | 2 | 0 |
| 2004 | 31 | 3 | 9 | 4 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 2 | 5 | 0 |
| 2007 | 34 | 10 | 20 | 5 | 0 | 0 | 0 | 1 | 0 | 0 | 3 | 0 | 0 | 1 | 4 | 0 | 2 | 0 | 2 | 2 | 0 |
| 2010 | 37 | 6 | 26 | 7 | 1 | 4 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 1 | 3 | 0 | 0 | 1 | 8 | 11 | 0 |
| 2010 | 37 | 10 | 31 | 8 | 0 | 5 | 0 | 6 | 0 | 0 | 2 | 0 | 0 | 5 | 7 | 1 | 0 | 4 | 5 | 7 | 0 |
| 2016 | 71 | — | 4 | 2 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 2 | 0 |
| 2016 | 82 | — | 7 | 2 | 0 | 3 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 |
| 2018 | 93 | — | 15 | 2 | 0 | 5 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 2 | 0 | 2 | 0 | 3 | 5 | 0 |
| 2018 | 95 | — | 7 | 2 | 0 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 3 | 0 |
| Total | | 12 | 167 | 44 | 1 | 29 | 0 | 17 | 3 | 1 | 18 | 1 | 0 | 17 | 28 | 5 | 7 | 6 | 29 | 49 | 0 |

Table S4.4. Random cluster sample without replacement from the *Journal of Field Archaeology* stratum. Key to column headings is in the left column of table S4.1.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Year** | **Vol** | **Iss** | **Articles** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** |
| 2000 | 27 | 1 | 4 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 2 | 2 | 2 | 0 |
| 2000 | 27 | 3 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 1 | 2 | 2 | 1 | 0 |
| 2003 | 28 | 3/4 | 13 | 4 | 0 | 2 | 0 | 0 | 1 | 0 | 4 | 1 | 0 | 4 | 0 | 2 | 0 | 2 | 4 | 3 | 0 |
| 2004 | 29 | 3/4 | 11 | 2 | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 3 | 3 | 0 |
| 2007 | 32 | 2 | 5 | 1 | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 2 | 1 | 0 |
| 2008 | 33 | 2 | 5 | 1 | 1 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12008 | 33 | 3 | 5 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 3 | 0 | 1 | 1 | 2 | 2 | 3 | 0 |
| 2009 | 34 | 4 | 6 | 2 | 1 | 2 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 2 | 1 | 0 |
| 2011 | 36 | 1 | 3 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 2 | 0 |
| 2012 | 37 | 2 | 4 | 2 | 0 | 2 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 2 | 0 |
| 2012 | 37 | 3 | 6 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 0 | 0 | 0 | 3 | 3 | 0 |
| 2015 | 40 | 4 | 6 | 1 | 0 | 2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 2 | 0 | 0 | 0 | 1 |
| 2015 | 40 | 6 | 5 | 3 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 2 | 1 | 0 |
| 2016 | 41 | 2 | 7 | 2 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| 2017 | 42 | 5 | 6 | 0 | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 0 |
| 2018 | 43 | 8 | 5 | 1 | 0 | 0 | 1 | 2 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 3 | 0 | 0 |
| 2019 | 44 | 6 | 4 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Total | | 17 | 103 | 22 | 2 | 21 | 2 | 9 | 7 | 0 | 7 | 3 | 0 | 26 | 4 | 5 | 8 | 8 | 29 | 23 | 1 |

Table S4.5. Random cluster sample without replacement from the *Journal of Archaeological Method and Theory* stratum. Key to column headings is in the left column of table S4.1.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Year** | **Vol** | **Iss** | **Articles** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** |
| 2000 | 7 | 4 | 4 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 |
| 2003 | 10 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| 2004 | 11 | 1 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2009 | 16 | 1 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2009 | 16 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2010 | 17 | 2 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2011 | 18 | 3 | 4 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 19 | 4 | 8 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 |
| 2014 | 21 | 2 | 12 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2015 | 22 | 1 | 9 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| 2015 | 22 | 4 | 10 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 3 | 1 | 0 |
| 2016 | 23 | 3 | 9 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 2016 | 23 | 4 | 12 | 2 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 0 |
| 2017 | 24 | 1 | 9 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 3 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2017 | 24 | 4 | 12 | 3 | 0 | 2 | 0 | 1 | 0 | 0 | 3 | 0 | 0 | 1 | 2 | 0 | 0 | 1 | 0 | 0 | 0 |
| 2018 | 25 | 4 | 6 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Total | | 16 | 108 | 14 | 0 | 6 | 0 | 3 | 3 | 0 | 13 | 0 | 2 | 8 | 6 | 1 | 4 | 1 | 5 | 5 | 0 |

# Evaluating the Stratification

Stratified sampling is only effective when there are genuine differences among the strata with respect to the variables of interest. In this instance, there are some reasons to suspect that the four journals reviewed differ somewhat in their emphases, with *Journal of Archaeological Method and Theory* having more emphasis on theory papers and less commonly dealing with sampling at all, *Journal of Field Archaeology* more likely to publish on field surveys or geoarchaeology, and *Journal of Archaeological Science* more likely to publish research that involves taphonomy, zooarchaeology, experiments, analytical chemistry, or isotopic studies.

Scanning the data in the tables above suggests that there are likely differences among the strata, with some markedly high or low values (in bold in table S4.1). Notably, *Journal of Archaeological Science* was considerably more likely to publish papers in which “sample” was used to describe individual specimens of bone, sediment volumes, and other kinds of scientific specimens and aliquots of material. *American Antiquity* was much less likely to include articles on experimental archaeology that failed to include randomization as a control, but more likely to use “carbon sample” to refer to individual dating specimens. *Journal of Field Archaeology* was much more likely than the others to have articles with unsubstantiated claims for “representative” samples.

Demonstrating these differences quantitatively might seem like a case for Analysis of Variance (ANOVA), but proportional data, such as we have here, violates its assumptions because it fits the binomial model (values between 0 and 1.0). An alternative for quick confirmation that there are differences among strata for the concerns that occurred reasonably frequently is to use the proportional estimates to provide expected frequencies for one-sample chi-square analysis of individual rows. Doing this with multiple rows would not be appropriate because the different sampling concerns are not independent (individual articles often scored positive on multiple concerns). This is not the ideal approach, but its exact outcomes are less important than demonstrating whether there are likely differences or not.

On several of the sampling concerns, these results do seem to confirm that the journals were reasonably distinctive strata, the null hypothesis being that the strata are all equivalent. For concern 1, failure to specify the type of sample, a result of (df = 3 in each case) exceeds the .05 threshold of 7.82. Concern 3 () and concern 8 () also exceed the .05 threshold. Even more easily, concerns 11 () and 16 () appear to differ significantly among journals. By contrast, on some concerns, journals did not seem to differ significantly, such as concern 5 (calling the whole population a sample, ).

A better alternative to chi-square is logistical regression, with the null hypothesis that there is no difference between the predicted proportions (from the entire stratified sample) and the actual proportions in the stratum samples. This also allows us to take into account the proportions for all concerns simultaneously, as long as they had no zeroes among their actual observations. Having transformed these into logit values (), I conducted a linear regression with the result of and F(1,46) = 39.9, with a probability of .5, indicating that the regression is weak and fit to the null hypothesis is not significant. A plot of the data shows only a loosely linear scatter of points with slope of *m* = .91, with several notable outliers that help to highlight differences. These include points for *American Antiquity* (concerns 1, 8, 11, 15), *JAS* (concerns 12, 17), *JFA* (concerns 6, 16), and *JAMT* (concerns 1, 5, 13, 15). Not surprisingly, this spread of outliers among the four journals is fairly consistent with the values bolded in table S4.1 and suggests pretty strongly that the strata do in fact differ from one another in multiple respects.

It is also possible to optimize the stratum sample sizes in disproportionate stratified sampling in terms of cost and expected variance (e.g., Deming 1950:213-238). For this paper, however, I was only concerned to ensure that sampling fractions were large enough to yield at least 100 articles per stratum. As the number of articles ranged from 103 to 167, this requirement was satisfied.

# References Cited

Deming, William Edwards

1950 *Some Theory of Sampling*. Dover Publications, New York.