

Appendix B: Rating preference across authors

We also analyzed whether there was any rating preference (or bias) across authors. In Figure A1, some authors in Snopes had a higher proportion of articles evaluating extreme ratings, such as *False* (e.g., Author F) or *True* (e.g., Author G), while others had a higher proportion of articles in the middle range, such as *Mixture* (e.g., Author H). Thus, in regard to ratings, it appears that authors at Snopes possess varying degrees of expertise relative to one another. However, in Logically, the ratings are relatively evenly distributed across authors compared to Snopes, indicating that different fact checkers may have different standards for the authors’ role in evaluating the accuracy of claims. Specifically, Logically employs Artificial Intelligence (AI) models to prioritize claims for debunking, which may have contributed to the relatively even distribution of ratings across authors.

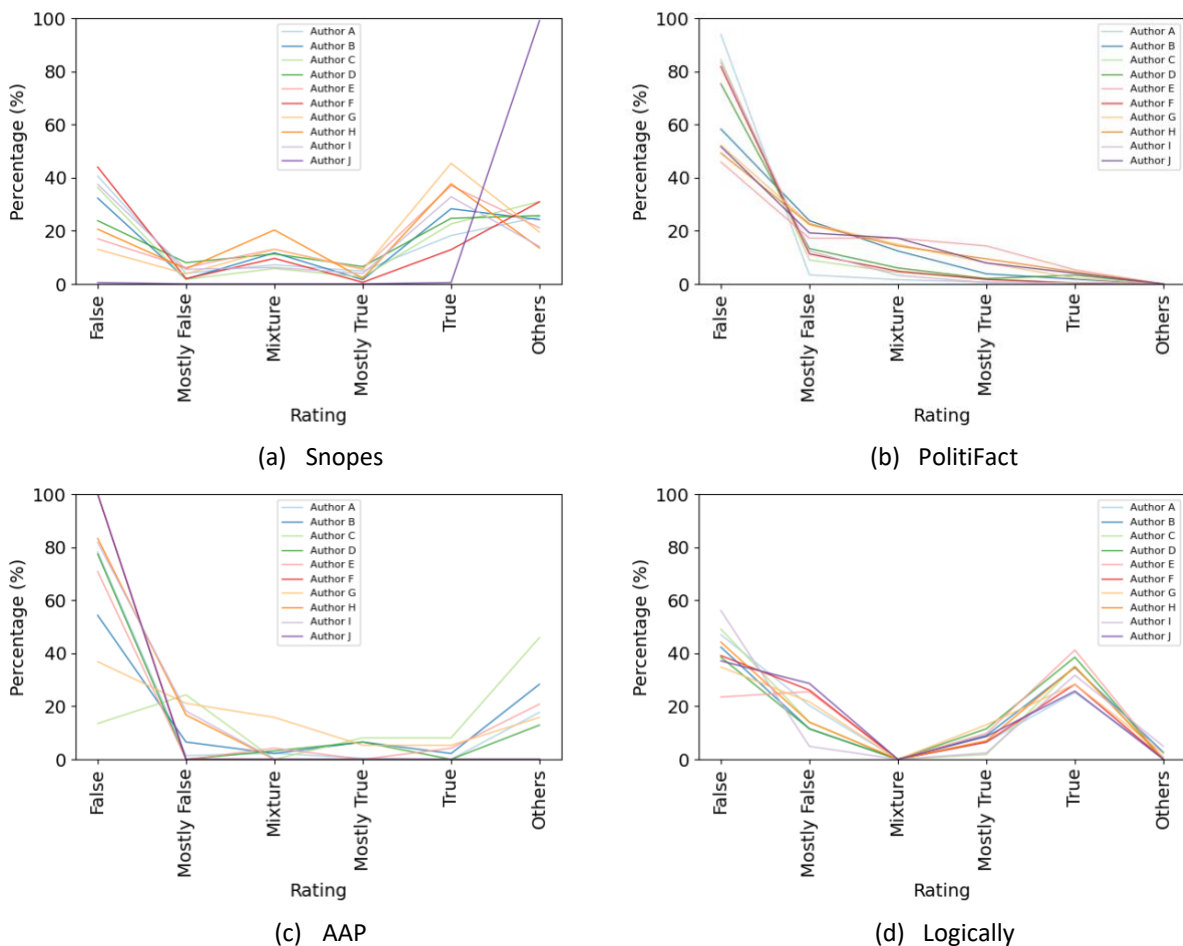


Figure A1. Comparison of rating distribution of fact-checking articles by authors across four fact checkers. Percentage shows the proportion of each rating among all fact-checking articles of each author.