methods

Jenine Harris

1/19/2022

[text on creating the portrait scale vars here?]

Given the lack of prior research on this topic, we will start with thorough exploratory data analysis and visualization to understand the sample characteristics, including examining bivariate descriptive statistics for each portrait variable and each GGH activity within religious affiliation groups. Our second step will be to build linear regression models with each GGH activity as the outcome and portrait variables as the predictors of interest. To select the portrait variables for regression modeling, we will conduct bivariate analyses examining the relationship between each portrait variable and each GGH activity. The group of portrait variables significantly associated with a moderate or stronger relationship with a GGH activity will be checked for multicollinearity and any portrait variables exceeding the VIF threshold of 2.5 will be dropped.

Following portrait variable selection, we will build a linear regression model for each of the six GGH activities with the following steps: (1) enter control variables (age, gender, education, income, health), (2) enter portrait variables and examine change in model fit (Model 2), (3) test whether trust in the healthcare system mediates the portrait variables (Model 3), (4) determine whether knowledge of genetics and rural/suburban/urban status moderate the relationship between the religious portrait variables and the GGH activities (Model 4). If mediation was confirmed in Model 2, Model 3 will test a moderated mediation model. If we do not confirm mediation in Model 2, trust in medicine will remain in the model as a covariate while we test a moderation model for Model 3. (5) For any significant moderators, estimate stratified models. We will test model assumptions (linearity, no perfect multicollinearity,homoscedasticity, and independent and normally distributed residuals) and conduct model diagnostics to identify any outlying or influential values.

For each model, model reporting will include overall model fit and significance along with coefficient values, direction, and significance. With a sample size of approximately 3,000 we will have more than 95% power ( = .05) to detect a small effect size in a linear regression model with up to 50 coefficients (Zhang and Mai 2021). Finally, we will conduct bivariate tests (e.g., Cramer’s V) of the portrait variables entered into each GGH activity model within religious affiliation groups to determine whether there are differences in which portrait variables are associated with GGH by religious affiliation.

Ensuring research results are reproducible and scientific information is freely available are two strategies to improve research quality (Harris et al. 2018) and reach . We will ensure our work is reproducible by including all relevant items from the CROSS survey checklist (Sharma et al. 2021) and all relevant methods detail described by Harris et al. (2018) in publications. We will also use publicly available software to conduct data management and analyses and by share the data and statistical code used in all dissemination in a free open access repository such as GitHub (Harris et al. 2018). All manuscripts published will be Gold open access to improve equity of access to our findings (Papin-Ramcharan and Dawe 2006).

scraps:

*Cluster analysis and description.* We will examine relationships among the portrait variables using scatterplots and correlation coefficients to gain an understanding of possible clusters containing portrait variables that are strongly correlated. We will use *fuzzy clustering*, a version of cluster analysis that indicates cluster membership along a 0 to 1 continuum rather than giving an absolute group assignment for each observation (Ferraro 2021).

*Cluster descriptions.* Following DuBois et al. (DuBois et al. 2019) we will examine the demographic characteristics and religious affiliation of participants in each cluster to determine any patterns.

DuBois, James M., Emily E. Anderson, John T. Chibnall, Jessica Mozersky, and Heidi A. Walsh. 2019. “Serious Ethical Violations in Medicine: A Statistical and Ethical Analysis of 280 Cases in the United States From 20082016.” *The American Journal of Bioethics* 19 (1): 16–34. <https://doi.org/10.1080/15265161.2018.1544305>.

Ferraro, Maria Brigida. 2021. “Fuzzy k-Means: History and Applications.” *Econometrics and Statistics*, November. <https://doi.org/10.1016/j.ecosta.2021.11.008>.

Harris, Jenine K., Kimberly J. Johnson, Bobbi J. Carothers, Todd B. Combs, Douglas A. Luke, and Xiaoyan Wang. 2018. “Use of Reproducible Research Practices in Public Health: A Survey of Public Health Analysts.” Edited by Conor Gilligan. *PLOS ONE* 13 (9): e0202447. <https://doi.org/10.1371/journal.pone.0202447>.

Papin-Ramcharan, Jennifer, and Richard A Dawe. 2006. “The Other Side of the Coin for Open Access Publishing–a Developing Country View.”

Sharma, Akash, Nguyen Tran Minh Duc, Tai Luu Lam Thang, Nguyen Hai Nam, Sze Jia Ng, Kirellos Said Abbas, Nguyen Tien Huy, et al. 2021. “A Consensus-Based Checklist for Reporting of Survey Studies (CROSS).” *Journal of General Internal Medicine*, 1–9.

Zhang, Zhiyong, and Yujiao Mai. 2021. *WebPower: Basic and Advanced Statistical Power Analysis*. <https://CRAN.R-project.org/package=WebPower>.