My aim as a methodologist is to address the disconnect between the characteristics of social science data and the requirements or availability of statistical methodology. I study topics related to missing data analysis, robust methods, and Structural Equation Modeling (SEM). On the topic of missing data, my research addresses the following research questions: **1) What mechanical processes lead to missing data and how does that impact statistical results? and 2) How can we improve the way we handle missing data?** These research questions are interesting and important because incorrect handling of missing data threatens the validity of statistical conclusions and correct handling of missing data relies on the reason for missingness. **My work on robust methods and SEM seeks to adapt or create robust statistical methodology that works for real, "messy" data.** This is interesting and important because many existing methods rely on assumptions that are not met for the data we commonly encounter in the social sciences, which may be non-normal, contain missing values, and/or have small sample sizes.

What mechanical processes lead to missing data and how does that impact statistical results? One thread of my work considers that participants can have missing values for different reasons, and thus correspond to different missing-data mechanisms – some participants overlook a survey question (MCAR) while others choose not to answer (MNAR). Although this is a likely scenario, common practice assumes that missing-data mechanisms are fixed across individuals. Gomer (2019, *Multivariate Behavioral Research*) was the first article to allow missingness to differ across individuals and evaluate the consequences for popular missing-data methods. Gomer & Yuan (2023b, *Multivariate Behavioral Research*) extends this work to the context of MNAR-based missing-data methods. Gomer (in preparation) evaluates the scenario in the context of confirmatory factor analysis and latent growth curve models while Kim & Gomer (in preparation) studies the topic in the context of multigroup invariance testing.

I have also named two subtypes of the missing not at random (MNAR) mechanism (*focused* and *diffuse* MNAR) which differ in the severity of their impact on statistical results in a foundational article that describes the features of these two subtypes in detail (Gomer & Yuan, 2021, *Psychological Methods*). These subtypes are relevant to all work involving the MNAR mechanism and are now included in the second edition of the premier textbook for missing data analysis (*Applied Missing Data Analysis* by Craig Enders, 2nd edition published in 2022).

How can we improve the way we handle missing data? I plan to pivot my missing data research to include more studies that address this research question, as this empowers researchers who are struggling with missing data. The latest thread of my research develops a novel framework for conducting sensitivity analysis with missing data. My work develops hypothesis tests and effect size measures that quantify how much statistical results are impacted by uncertainty in the missing-data mechanism (Gomer, in preparation). In the future, I plan to develop a novel multiple imputation approach for ordinal data with non-ignorable missingness. This is motivated by a hypothetical scenario in which MNAR missingness occurs on ordinal items with a 4-point Likert scale – in other words, a scenario in which item responses cannot be treated as continuous and it's desirable to handle MNAR missingness at the item-level. I also plan to collaborate with Jon Butner to identify a suitable multiple

imputation procedure in the context of transformations in dynamical systems.

How can we adapt or create robust statistical methodology that works for real, "messy" data? My past work includes developing robust methodology applicable to equivalence testing in mean and covariance structure analysis (Yuan, Gomer, & Marcoulides, 2021, *Multivariate Behavioral Research*). On the topic of SEM, I developed new effect size measures that quantify model misspecification (Gomer et al., 2018, *Structural Equation Modeling*). A key concern in SEM is that of model fit and model misspecification; i.e. how well/badly the hypothesized relationships are supported by the data. The effect size measure with the best performance, $\mathscr{E}$, captured misspecification and was not undesirably impacted by sample size, distribution, or model size, so it is designed to work well for "messy" characteristics of real data. Gomer et al. (in preparation) extends this work by developing confidence intervals for $\mathscr{E}$ using bootstrap methodology.

Other topics: I also have a few projects which don't neatly fall into the categories of my research questions. For example, I investigated an alternative framework to improve the generalizability of Monte Carlo simulation studies (Gomer, Lee, & Kim, in draft). As I have encountered in my own experience, a strong pattern of results may appear to be stable but fail to generalize with a seemingly insignificant change to the population model. However, it is rarely feasible to consider a large number of population models and run 500 or 1000 replications each. Instead, my proposed approach breaks up 1000 replications of a given simulation condition into smaller subsets, each using a different population model. The results on overall bias and efficiency are promising and suggest the approach may be a viable strategy. In the future, I plan to develop a survey instrument to measure academic contrapower harrasment. Contrapower harassment occurs when someone in a lower status position harasses someone in a higher status position; in the context of academia, this could be a student harassing a professor over grades with behaviors like threats and verbal abuse. Currently, there is no survey instrument that measures academic contrapower harassment but it is important to measure because it can disproportionately affect certain populations and lead to highly trained professionals leaving academia in favor of industry.

Substantive Interests: Substantively, I enjoy applying statistical methodology to address research questions in psychology and other disciplines. Some of my past and current collaborations with substantive colleagues include topics such as: 1) the mediation of coping strategies on the relationship between comorbidities and emotional quality of life in cancer patients; 2) the social context of non-suicidal self-injury and how it relates to severity of NSSI behaviors, suicide risk, and other social factors; 3) the cross-cultural equivalence of a scale used to assess emotion regulation ("Responses to Positive Affect") in Korean and US samples; 4) novel technology for research on vaping involving rodents, called the Rodent Electronic Nicotine Delivery System;and 5) health disparities in Latinx mental health.