Mixed design data is a mixture of paired and independent samples. For example, data of this type can arise in matched-pair designs where, due to some circumstance, parts of the sample may only have responses on one element of a pair. Hence, there are complete data obtained from \((n_0)\) matched pairs and partial data from only the first element (say \(n_1\) observations) or second element (say \(n_2\) observations) of a configured pair. In the univariate case, several test procedures exist that can be used to analyze this type of data. However, in the multivariate setting, there are only a few existing approaches.

Some of these approaches are rather “naïve” in the sense that they simply utilize only the paired or partial data. For instance, these “naïve” approaches include disregarding the partial data and carrying out a paired sample test on the remaining paired data \((n_0)\), disregarding the paired samples and carrying out a two independent sample test on the remaining partial samples \((n_1\) and \(n_2)\), and ignoring the correlation among the paired data and treating the data as if it came from two independent samples \((n_0+n_1\) and \(n_0+n_2)\). On the other hand, an approach that utilizes all of the data is the Wald type mean-based procedure developed by Adkison-Johnson et al. (in press). However, since it is known that
the mean is sensitive to the presence of outliers, an alternative procedure based on sample medians, which is known to have desirable robustness properties, is proposed in this study.

More specifically, this dissertation proposes the use of component-wise medians for the purpose of estimation and inference in this multivariate mixed design context. Asymptotic linearity and quadraticity results are derived and these results lead to the joint asymptotic distribution of the estimates. Subsequently, this result yields a Wald type median-based procedure analogous to that of by Adkison-Johnson et al. (in press). The proposed methodology is illustrated on some examples and compared against the mean-based procedure, as well as some of the “naive” approaches, through a simulation study. Results indicate that the median-based procedure is superior to the “naive” approaches for any proportion of paired observations in the sample. Additionally, it is superior to the mean-based procedure under heavy tailed distributions with at least 10% data contamination. Also, for distributions that slightly deviate from a multivariate normal distribution, the median-based procedure is almost as good as the mean-based procedure.