S&DS 363 Final Project:
A Multivariate Exploration of Ace-Aro Relational Experiences

May 7, 2022

Introduction: Ace-Aro Identities & Relational Health

The asexual- and aromantic spectrum (hereafter referred to by the shorthand term "ace-aro") is an emerging yet crucial identity community within the LGBTQ+ umbrella. Referring, respectively, to individuals who lack sexual or romantic attraction, asexuality and aromanticism are identities that fundamentally challenge *compulsory sexuality*. This ubiquitous sociocultural assumption, analogous to the more commonly known term *compulsory heterosexuality*, posits that all people experience sexual attraction and desire (Gupta, 2015). It is norms like compulsory sexuality and compulsory heterosexuality that fuel *amatonormativity*, the societal ascription of value and superiority to (exclusive) sexual and romantic relationships, and the myriad norms and implications that result from this (Vares, 2021). While all ace-aro people share some commonality of experience surrounding amatonormativity and its impacts, this community is by no means a monolith. Both asexuality and aromanticism exist as spectra: the degree to which an ace-aro person may desire sexual and/or romantic intimacy can vary, and it is also possible for a person to be asexual but *not* aromantic, or vice versa. In all, there is a rich diversity in identity and experience along the ace-aro spectrum, yet this LGBTQ+ community remains critically understudied in both queer studies and in public health. This final project builds upon the work done by one of the group members (who identifies as an ace-aro individual themselves) for their senior thesis in WGSS, which examined ace-aro relational health through an original mixed-methods public health study.

In the LGBTQ+ public health field, scholars have made various attempts to study ace-aro identity formation, relationships, and health from both quantitative and qualitative angles. Over

the past decade, several authors, particularly those interested in ace-aro relationships, have found empirical evidence indicating that this community may be at disproportionate risk for relational violence.[1] The pervasive norms and assumptions resulting from amatonormativity can make ace-aro people "feel obliged to engage in [sexual and/or romantic activity]" against their wishes, and the frequency with which coercive and potentially traumatic relational experiences may be happening in this community is alarming (Carrigan, 2011; Prause & Graham, 2007; Robbins et al., 2016). Countless firsthand accounts in the published literature, from interviews and conversations with ace-aro individuals, detail experiences of coercion under assumptions of normalcy. For instance, many ace-aro people express partaking in certain sexual and/or romantic activities purely because they feared hostility from their partner(s) otherwise or because they assumed such behavior was necessary to a relationship. Intense emotional stress and long-lasting trauma can result from such experiences, and in other LGBTQ+ populations, similar phenomena have already been recognized and studied as public health concerns. The same is not yet true of the ace-aro community, despite the unique vulnerabilities it faces. Our group member's senior thesis intervened within this space in the literature, and by expanding upon it with multivariate analyses, we are continuing to add meaningfully to ace-aro scholarship and studies on sexual harm more broadly. Ultimately, deconstructing compulsory sexuality, amatonormativity, and their negative impacts is a worthwhile project for *all* communities, as projects like this one invite us to reckon with our understandings of consent and harm in important ways.

[1] Throughout this report, the intentionally broad terms "relational violence" and "relational trauma" are used frequently. We use these terms to encompass any and all forms of harm which can occur through sexual and/or romantic relationships, including but not limited to sexual trauma, intimate partner violence (IPV), physical violence, and emotional abuse.

Research Questions & Project Design

The thesis project we are building from included some quantitative analyses, but they were exclusively univariate — in this extension, we hope to use multivariate techniques to add to the insights gathered in the original project. The survey had three main outcome variables of interest: sexual trauma, intimate partner violence (IPV), and identity-based emotional abuse. Multivariate analyses allow us to gain a richer understanding of ace-aro relational trauma by assessing these three separate measures of it at once, rather than separately and independently. We will be utilizing four different multivariate techniques throughout this project: multivariate analysis of variance (MANOVA) and generalized linear models (GLMs), factor analysis (FA), and discriminant analysis (DA).

Our main research questions are as follows: (1) What are the main drivers of overall vulnerability to relational trauma in the ace-aro community? (2) Are certain identities or subgroups within the ace-aro community at higher risk of relational trauma than others? We hope to use MANOVA and FA to answer the first question and DA along with multivariate GLMs to answer the second. To study the first question, we will assess the significance of various predictors for the three studied forms of relational trauma: sexual trauma, IPV, and emotional abuse. GLMs will allow us to analyze linear relationships between predictors and response variables, whereas FA will help uncover what underlying constructs, if any, may be linked to experiences of trauma. Next, we will use DA and MANOVA to distinguish the risk of relational trauma between subgroups based on sexual orientation and/or gender. The thesis found some preliminary evidence for disparities in risk between ace-aro individuals (i.e., those identifying explicitly as “asexual” and/or “aromantic”) compared to those identifying with other labels on

the ace-aro spectrum, such as demisexual or gray-asexual. Our analysis might be able to add further depth to this, and expand on other important findings of the thesis.

Dataset Basics & Summary Statistics

We are using a cleaned and de-identified version of our group member's senior thesis dataset. The data was collected between December 30-31, 2021, and all participants were recruited via posts made in ace-aro community pages on Tumblr and Reddit. The data was gathered through an online survey. A total of 745 individuals completed the survey, and for each person the survey captured data on demographics, relational trauma, and various identity-related characteristics. The variables, as they appear in our dataset (the cleaned version used in this project), are listed and described below:

- **race** (categorical)
- **ethnicity** (categorical)
- **education** (categorical)
- **gender** (categorical)
- **sexual_orientation** (categorical)
- **acceptance_concern** (continuous, "Acceptance Concerns")[2]
- **conceal_motiv** (continuous, "Concealment Motivation")
- **id_uncertain** (continuous, "Identity Uncertainty")
- **internalized_stigma** (continuous, "Internalized Stigma")
- **difficult_process** (continuous, "Difficulty Processing")
- **id_superior** (continuous, "Identity Superiority")
- **id_aff** (continuous, "Identity Affirmation")
- **id_central** (continuous, "Identity Centrality")
- **ses_17** (binary indicator for sexual trauma)[3]
- **ses_score** (continuous, frequency/severity of sexual trauma)
- **cts_score** (continuous, frequency/severity of IPV)
- **ia_score** (continuous, frequency/severity of identity-based emotional abuse)

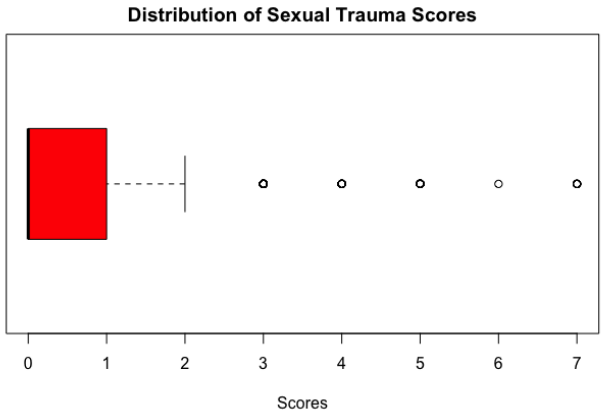
[2] Eight identity-related scores were calculated through the LGB Identity Scale (LGBIS), a public health survey measure that has been used frequently in LGBTQ+ health studies. All eight are listed here and were used in the thesis study.

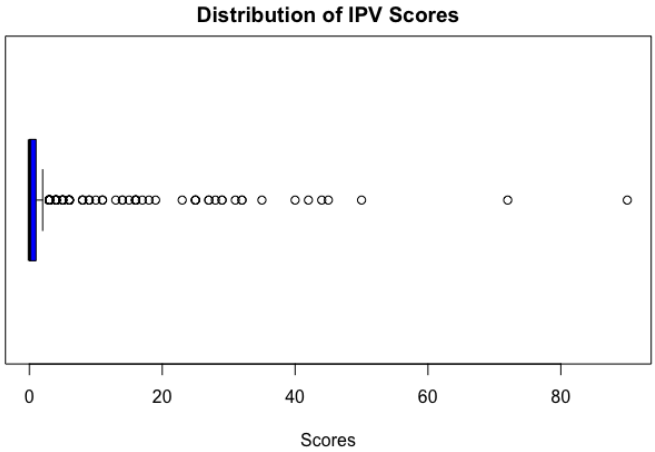
[3] The survey measure used to capture this is called the Sexual Experiences Survey, or "SES" for short. This is why variables related to sexual trauma use the acronym SES. Similarly, data on IPV was collected through the Conflicts Tactics Scale (CTS) and identity-based emotional abuse data was gathered through the Identity Abuse Measure (IA). The SES consisted of seven questions, each pertaining to a specific adverse sexual eperience. ses_17 is recorded as 1 for an individual who reported having had at least one of these experiences, and 0 otherwise.

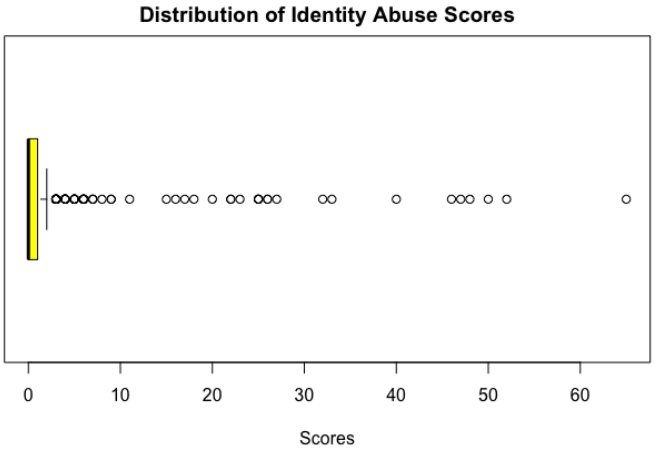
While most of these variables — especially those capturing information on relational trauma and identity-related predictors — are continuous and calculated using statistically validated public health measures, this is by no means a perfect data set. It is severely lacking in racial and ethnic diversity. Additionally, it is difficult to capture data on sexual coercion; most survey measures in existing public health literature are designed to quantify non-consensual sexual experiences, such as assault or rape, but lack the nuance needed to represent coercion or implicit pressure from a partner accurately. Despite these shortcomings, the dataset is still unique in the information it holds on the ace-aro population and can provide many important insights.

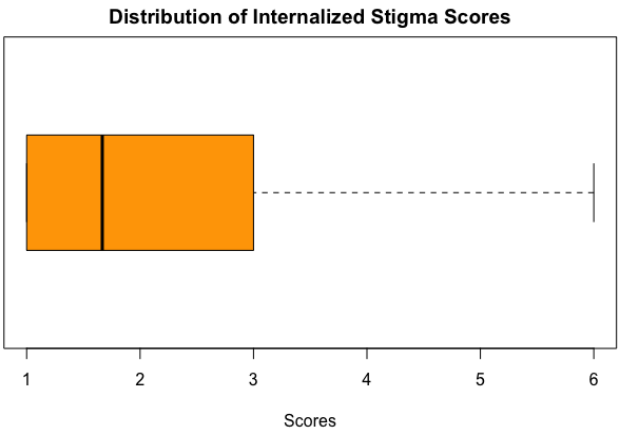
Descriptive Plots for Key Variables

We anticipate that a few identity-related variables may be significant predictors of sexual harm and vulnerability. To get an initial sense of what the distribution of these factors may look like, we generated a series of five boxplots, representing the following key variables: sexual trauma, IPV, identity abuse, internalized stigma, and identity affirmation.







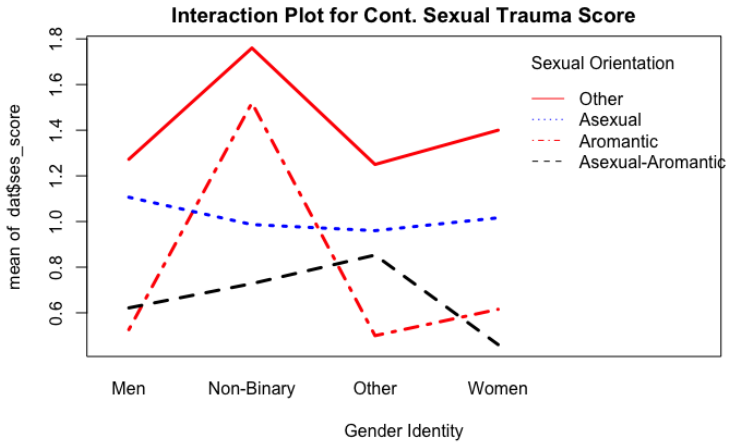


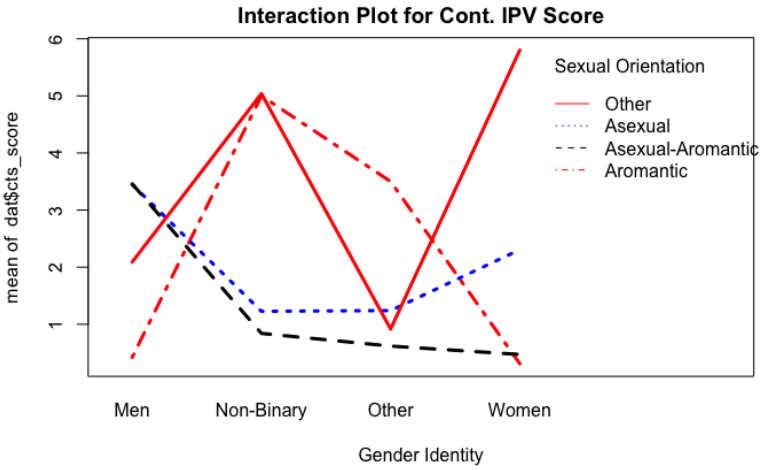
These figures show skewed distributions for the three outcome variables, necessitating the need for some kind of transformation to ensure multivariate and univariate normality. In DA and FA, we used log-transformed versions of these scores for this reason.

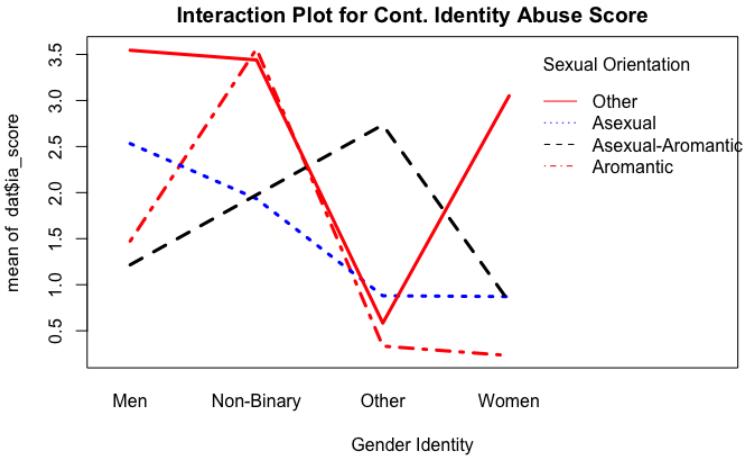
An additional and important observation here is that identity-related factors, such as internalized stigma and identity affirmation, all have minimums of 1 and maximums of 6. It is interesting (and encouraging) to note, though, how much higher the mean score for identity affirmation is than that for internalized stigma. With this background information on our project and dataset established, we now move into our multivariate analyses.

Multivariate Technique #1: MANOVA and Multivariate GLMs

This section will first use MANOVA to assess the significance of group differences across sexual orientation and gender identity, before using linear modeling to more explicitly study the predictors of relational violence. Here, we are interested in two main categorical variables — sexual orientation (Asexual, Aromantic, Ace-Aro, or Other) and gender identity (Men, Women, Non-Binary, or Other) — as well as eight continuous predictors, which are the LGBIS subscales. These continuous predictors include internalized stigma, identity affirmation, identity uncertainty, and other related constructs, all scored between 1 and 6. We begin the MANOVA analysis process by analyzing interaction plots for our categorical predictors:







Interaction plots allow us to assess whether our response variable(s) may be impacted by, in this case, gender identity and/or sexual orientation. Observing intersecting lines on an interaction plot would suggest the presence of another, non-additive factor that impacts the response variable(s). We do see such intersections on all three of these plots, implying that for all three response variables, there is a non-additive interaction effect between sexual orientation and gender identity. Practically speaking, this means that we should add interaction terms into our multivariate GLMs and be cognizant of these effects when conducting MANOVA.

With this in mind, we went on to fit two-way MANOVA on our data and obtained both univariate and multivariate results (only the latter are included/discussed below). Our model was fit to predict sexual trauma, IPV, and identity abuse using an interaction term between sexual orientation and gender; ultimately, only the intercept and sexual orientation were significant based on *all four* multivariate test statistics. As shown below, though, the other predictors

(gender and the interaction term) were significant based only on Roy's Largest Root. We overall concluded from this that we can be more certain of differences in group means across sexual orientations rather than across gender identities.

```
Term: (Intercept)

Sum of squares and products for the hypothesis:
          ses_score cts_score  ia_score
ses_score  416.3316  980.4667  779.7461
cts_score  980.4667 2309.0126 1836.3129
ia_score   779.7461 1836.3129 1460.3839

Multivariate Tests: (Intercept)
                 Df test stat approx F num Df den Df     Pr(>F)
Pillai            1 0.2128102 65.51278      3    727 < 2.22e-16 ***
Wilks             1 0.7871898 65.51278      3    727 < 2.22e-16 ***
Hotelling-Lawley  1 0.2703416 65.51278      3    727 < 2.22e-16 ***
Roy               1 0.2703416 65.51278      3    727 < 2.22e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

------------------------------------------

Term: sexual_orientation

Sum of squares and products for the hypothesis:
          ses_score cts_score ia_score
ses_score  33.63023  82.20093 33.52374
cts_score  82.20093 243.45332 86.01064
ia_score   33.52374  86.01064 72.78848

Multivariate Tests: sexual_orientation
                 Df test stat approx F num Df   den Df    Pr(>F)
Pillai            3 0.0240650 1.965028      9 2187.000 0.0395578 *
Wilks             3 0.9760051 1.971879      9 1769.478 0.0389021 *
Hotelling-Lawley  3 0.0245131 1.976480      9 2177.000 0.0382563 *
Roy               3 0.0212558 5.165161      3  729.000 0.0015454 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

------------------------------------------
```

```
Term: gender

Sum of squares and products for the hypothesis:
          ses_score cts_score  ia_score
ses_score  14.98530   36.4662  52.23167
cts_score  36.46620  121.5377 152.65709
ia_score   52.23167  152.6571 243.52571

Multivariate Tests: gender
                 Df test stat approx F num Df   den Df   Pr(>F)
Pillai            3 0.0161835 1.317974      9 2187.000 0.221978
Wilks             3 0.9838520 1.319564      9 1769.478 0.221330
Hotelling-Lawley  3 0.0163769 1.320463      9 2177.000 0.220687
Roy               3 0.0138756 3.371777      3  729.000 0.018124 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

------------------------------------------

Term: sexual_orientation:gender

Sum of squares and products for the hypothesis:
          ses_score cts_score  ia_score
ses_score  19.13264  88.80663  53.13861
cts_score  88.80663 985.34418 234.07919
ia_score   53.13861 234.07919 276.33380

Multivariate Tests: sexual_orientation:gender
                 Df test stat approx F num Df   den Df   Pr(>F)
Pillai            9 0.0447039 1.225264     27 2187.000 0.196357
Wilks             9 0.9558068 1.226876     27 2123.857 0.195000
Hotelling-Lawley  9 0.0457036 1.228355     27 2177.000 0.193658
Roy               9 0.0288902 2.340110      9  729.000 0.013278 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Next, we conducted a linear hypothesis test using a "combined" factor variable for sexual orientation and gender. Based on our two-way MANOVA results, we were interested in further assessing mean differences across sexual orientation groups, and we could do this using contrasts. We made a multivariate comparison between the "Ace-Aro" and "Other" groups, and all four multivariate test statistics were significant for this comparison: this provides rather strong evidence for a difference in multivariate means between these two identity groups, and the univariate hypothesis test found a consistent result. These conclusions — that there likely is some kind of difference in relational violence experiences between these groups — also account

for gender identity since we used the combined factor variable in our testing process. Our output for these analyses is included below:

```
Sum of squares and products for the hypothesis:
          ses_score cts_score  ia_score
ses_score  28.60974  105.6168  33.25204
cts_score 105.61677  389.8987 122.75445
ia_score   33.25204  122.7545  38.64762

Sum of squares and products for error:
          ses_score cts_score  ia_score
ses_score  1623.668  2945.366  1591.799
cts_score  2945.366 35214.965 12427.293
ia_score   1591.799 12427.293 26507.437

Multivariate Tests:
                 Df test stat approx F num Df den Df    Pr(>F)
Pillai            1 0.0207837 5.143495      3    727 0.0015927 **
Wilks             1 0.9792163 5.143495      3    727 0.0015927 **
Hotelling-Lawley  1 0.0212249 5.143495      3    727 0.0015927 **
Roy               1 0.0212249 5.143495      3    727 0.0015927 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
Linear hypothesis test

Hypothesis:
so_gender_combAsexual - AromanticNon - Binary  + so_gender_combAsexual -
AromanticOther  + so_gender_combAsexual - AromanticWomen -
so_gender_combOtherMen - so_gender_combOtherNon - Binary -
so_gender_combOtherOther - so_gender_combOtherWomen = 0

Model 1: restricted model
Model 2: ses_score ~ so_gender_comb

  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1    730 1652.3
2    729 1623.7  1     28.61 12.845 0.0003609 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

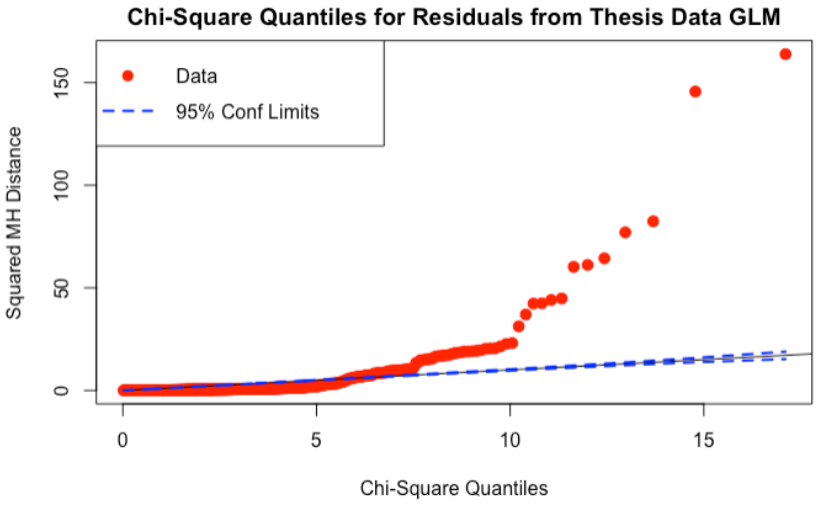
While our discussion thus far has centered largely on our second research question — the intra-group differences that may exist in relational violence experiences — we also wanted to use GLM techniques to examine our first question. We fit a multivariate GLM to assess the significance of certain categorical *and* continuous predictors on relational violence. In addition to a sexual orientation/gender interaction term, we added internalized stigma, difficulty processing, and identity affirmation nto the model. Our multivariate response variable consisted of sexual trauma, IPV, and identity abuse. Interestingly, we found that the three continuous predictors were statistically significant under all four multivariate test statistics (relevant output included below). The categorical predictors, including the interaction term, were only significant based on Roy’s Largest Root. This was insightful, as it implies that sexual orientation, for instance, may not be as indicative of an ace-aro individual’s vulnerability to relational harm as we might think. Whereas MANOVA suggested that significant differences in types or frequency of relational violence might exist between sexual orientation groups, our GLM indicates that membership in

one of these groups does not necessarily increase or decrease the likelihood of adverse experiences. Instead, other identity-related constructs may matter more.

```
Term: internalized_stigma

Sum of squares and products for the hypothesis:
           ses_score   cts_score   ia_score
ses_score  0.3479600   0.4743072  -9.053674
cts_score  0.4743072   0.6465322 -12.341140
ia_score  -9.0536737 -12.3411405 235.570233

Multivariate Tests: internalized_stigma
                 Df test stat approx F num Df den Df   Pr(>F)
Pillai            1 0.0116896 2.854459      3    724 0.036419 *
Wilks             1 0.9883104 2.854459      3    724 0.036419 *
Hotelling-Lawley  1 0.0118279 2.854459      3    724 0.036419 *
Roy               1 0.0118279 2.854459      3    724 0.036419 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

----------------------------------------

Term: id_aff

Sum of squares and products for the hypothesis:
           ses_score  cts_score  ia_score
ses_score  10.643204   8.082634 -22.87021
cts_score   8.082634   6.138092 -17.36803
ia_score  -22.870207 -17.368032  49.14369

Multivariate Tests: id_aff
                 Df test stat approx F num Df den Df   Pr(>F)
Pillai            1 0.0108338 2.643197      3    724 0.048299 *
Wilks             1 0.9891662 2.643197      3    724 0.048299 *
Hotelling-Lawley  1 0.0109525 2.643197      3    724 0.048299 *
Roy               1 0.0109525 2.643197      3    724 0.048299 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

----------------------------------------
```

```
Term: difficult_process

Sum of squares and products for the hypothesis:
           ses_score cts_score   ia_score
ses_score 14.728307  48.09567 -2.9141490
cts_score 48.095673 157.05769 -9.5162302
ia_score  -2.914149  -9.51623  0.5765947

Multivariate Tests: difficult_process
                 Df test stat approx F num Df den Df   Pr(>F)
Pillai            1 0.0120250 2.937355      3    724 0.032585 *
Wilks             1 0.9879750 2.937355      3    724 0.032585 *
Hotelling-Lawley  1 0.0121714 2.937355      3    724 0.032585 *
Roy               1 0.0121714 2.937355      3    724 0.032585 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

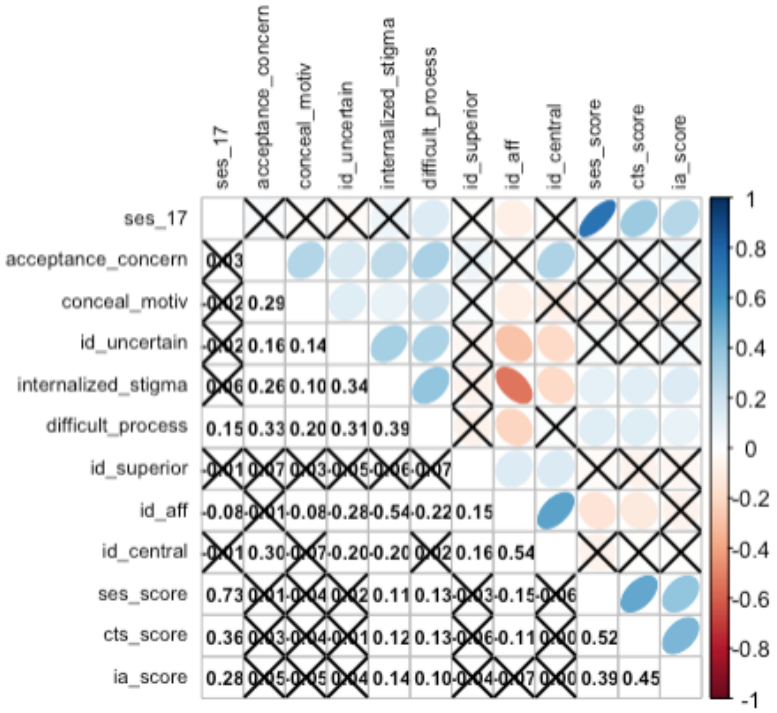
Finally, before concluding our discussion on MANOVA and GLM techniques, we want to check our overall model assumptions and make concluding comments on the interpretability of our results. We generated a chi-square quantile plot of our model's residuals, and if our assumptions were met, we would assume this plot to look roughly linear. Unfortunately, as we can see on the right, this was not the case. Given the nature of this dataset and how



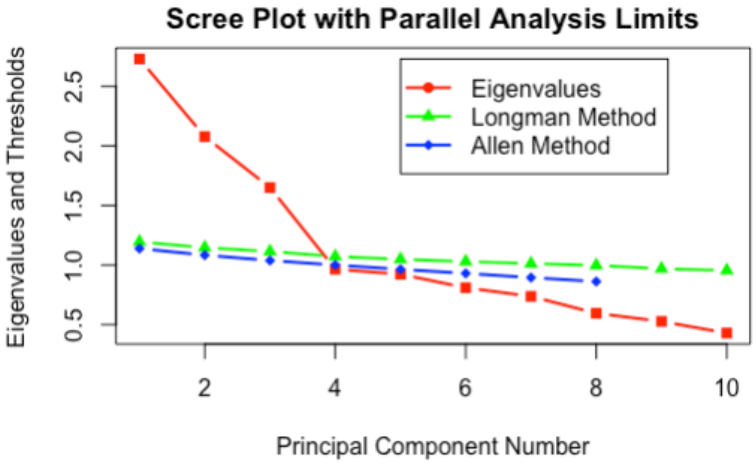
certain variables — such as CTS and IA scores — were calculated, we did not necessarily expect that it would be a perfect fit for GLM-fitting. This residual plot confirms this and means that we must take our results critically. This does not necessarily, however, invalidate some of the more general trends or associations we observed. For instance, we still acknowledge the existence of group mean differences across sexual orientations, and it is entirely possible that the effect of identity-related continuous variables on the likelihood of relational violence is significant.

Multivariate Technique #2: Factor Analysis

Since our conclusions from MANOVA and multivariate GLM analyses may be somewhat unreliable due largely to difficulties in meeting the assumptions of these techniques, we continued to assess our research questions with Factor Analysis. FA is informative towards our first research question: what drives heightened vulnerability to relational harm in the ace-aro community? By identifying key underlying "factors" and developing a related story around our dataset, we hoped to understand better which variables were associated with relational trauma and why. Before beginning FA, we made a few variable adjustments, especially based on how our residual plots looked in the first section. Namely, we log-transformed CTS and IA scores before making our correlation matrix. A correlation plot is shown at left charting the relationships between each variable, with colored ovals denoting the



direction and magnitude of statistically significant correlations. There are not too many strong (large-magnitude) correlations present, though some exist. There is a strong negative correlation between identity affirmation and internalized stigma, for instance, as well as a positive one between identity centrality and identity affirmation. Additionally, some of the outcome variables are strongly correlated, such as sexual trauma and IPV. These are overall rather intuitive results but do not definitively suggest that FA will work well with our data. To better determine that, we used the r package rela to calculate the KMO value of the dataset, which turned out to be 0.66; unfortunately, this statistic *does not* indicate that this dataset is a strong fit for factor analysis. Nonetheless, as with MANOVA, we have decided to proceed in our analyses with the understanding that our results must be taken critically. From here, our first step was to determine how many factors we anticipate may exist. We did this by running principal components analysis and generating both a parallel plot[4] and a scree plot. For brevity, we have included only our parallel analysis results below:



Our scree plot and parallel analysis both indicate that using either 3 or 4 factors may be appropriate; in the analyses that follow, we decide to use 3. The next major decision we had to make was regarding our extraction method. In our code, we conducted FA using both maximum likelihood (ML) and principal axis factoring (PAF) extraction methods and ultimately found that ML worked

[4] We were able to do this post-log transformations, as those operations gave our data a somewhat more multivariate normal distribution (although admittedly still imperfect).

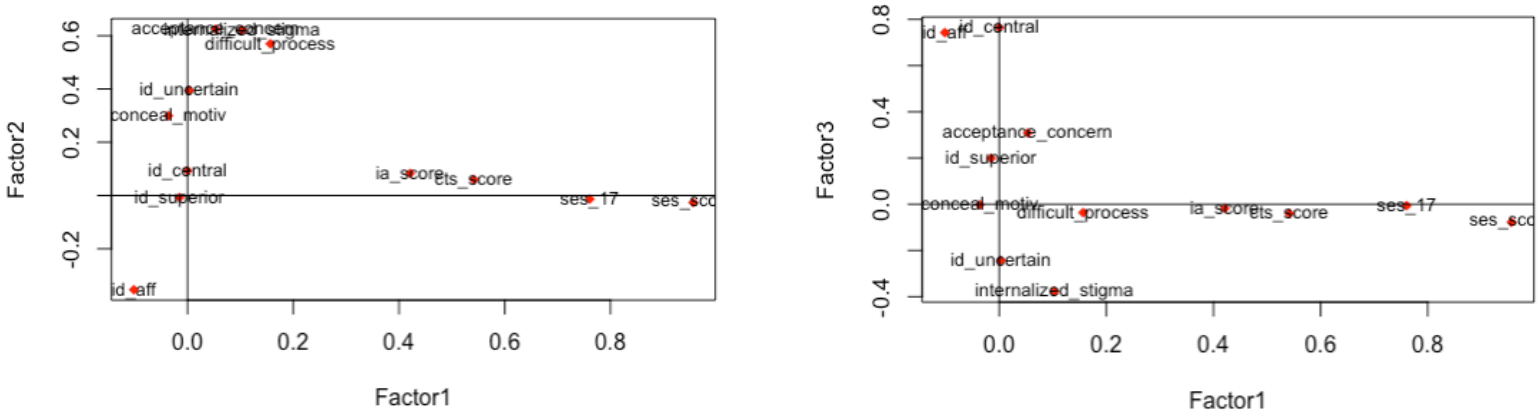
better. To keep this report concise, we will only be displaying and discussing our results which used ML (again, with three factors).

Based on the loadings we observed for each variable (relevant output is included at right), across our three factors, we were able to draw some conclusions regarding how the factors might be interpreted. For instance, the outcome variables all carried very high loadings on Factor 1, which implies that they are related by virtue of representing relational trauma. What is interesting about this is that it suggests that much of the relational trauma experienced by ace-aro individuals may occur at the hands of intimate partners (since relational violence more broadly appears to be closely associated with IPV). Factor 2 was very different and featured no significant loadings on any of the outcome variables. Instead, negative identity-related experiences, such as acceptance concerns and internalized stigma, seemed to play a bigger role. In this sense, Factors 2 and 3 are somewhat similar: they both feature high loadings on identity-related constructs instead of outcomes, but while Factor 2 is more negative, Factor 3 is associated with positive identity experiences (such as identity affirmation and identity centrality).

```
Loadings:
                   Factor1 Factor2 Factor3
ses_17              0.761
acceptance_concern          0.625   0.309
conceal_motiv               0.300
id_uncertain                0.395  -0.244
internalized_stigma 0.103   0.621  -0.378
difficult_process   0.156   0.570
id_superior                         0.199
id_aff             -0.102  -0.354   0.743
id_central                          0.765
ses_score           0.956
cts_score           0.541
ia_score            0.421
```

While we tried out different options in our code, we ultimately settled on using a varimax rotation in this analysis, and the resultant loading plots are included below. Overall, the axes pertaining to each factor are clear and easy to interpret, especially given the loadings discussed above. In all, our FA results are not as helpful in answering our research question as we had initially hoped: we expected that some of the outcome variables might be associated, via some

underlying factor(s), with key identity-related constructs, but this isn't necessarily the case as the outcome variables are largely associated *with each other*.

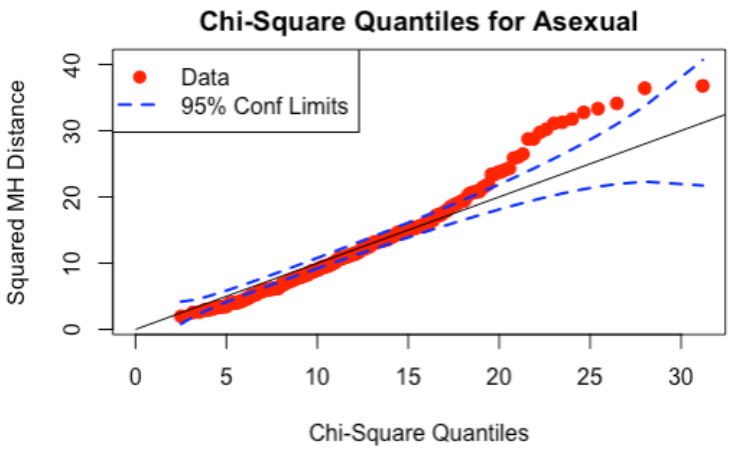


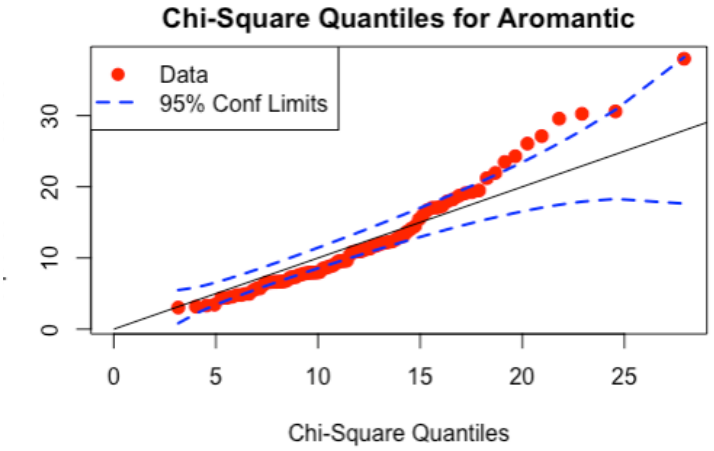
As a final "check" on our FA analysis, we generated a reproduced correlation matrix and used it to calculate (1) root mean squared residuals (RMSR) and (2) the percentage of "large" residuals. The latter is perhaps most telling: 11% of our residuals were greater than 0.05, which is not a superb indication that FA worked well. While our particular selections — choosing to use ML as our extraction method, for instance, and adding a varimax rotation — worked better than the alternatives we tried (PAF, for instance, resulted in 15% large residuals), our analysis was still imperfect. Indeed, this was the result we expected upfront based on our KMO calculation. Thus, as with MANOVA, our results in this section must be interpreted with caution.

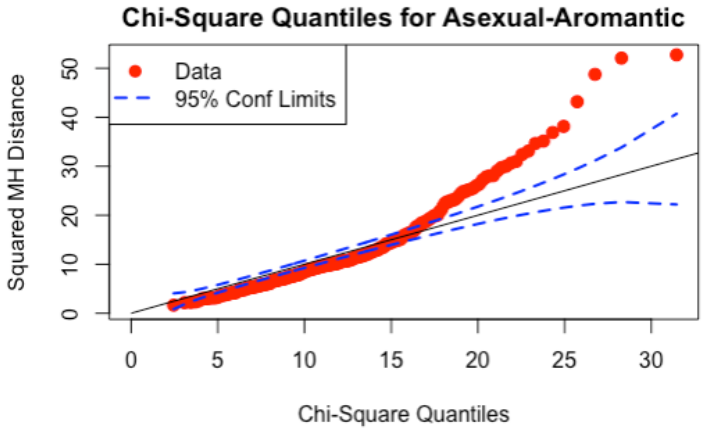
Multivariate Technique #3: Discriminant Analysis

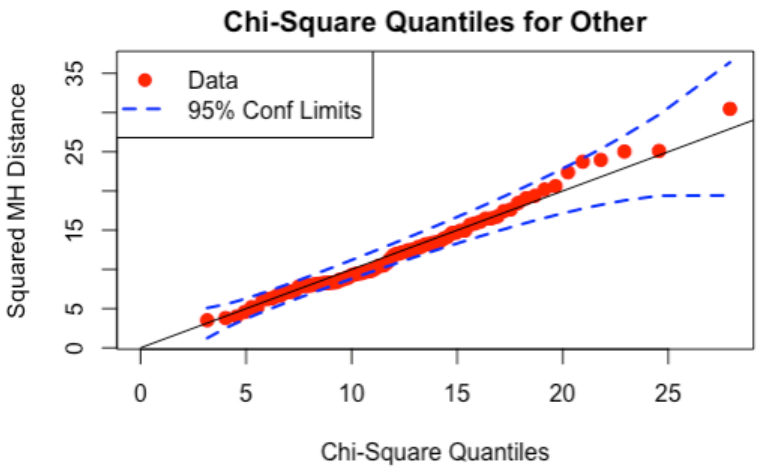
To further explore our second research question about intra-community differences, we used DA to examine which variables, if any, may help us distinguish meaningfully between sexual orientation groups. For instance, as MANOVA revealed, it is possible that individuals of other ace-aro spectrum identities may experience relational violence to a significantly different

degree than their counterparts who identify exclusively as "Ace-Aro." Additionally, we conducted a Wilks' Lambda test and found further evidence for such differences (the test resulted in a p-value much smaller than 0.05). Before formally beginning DA, we created chi-square quantile plots for each of our four sexual orientation groups to check for multivariate normality. Based on our last two sets of analyses, we decided to log-transform all three of our outcome variables before doing this, and it does appear that this helped somewhat:









While the first three of these plots certainly do not look perfect, they are much closer to representing multivariate normality than pre-transformations. Knowing, again, that we may need

to interpret our DA results with some caution, we proceeded with our analysis. We began by conducting a Box's M test for differences in covariance matrices; in DA, we want covariance matrices to be roughly equivalent and thus fail to reject the null hypothesis. Unfortunately, our p-value was sufficiently small in our test that this did not occur. This is yet another indication that we will need to interpret our results critically.

```
[1] "Raw (Unstandardized) Coefficients"
                      LD1   LD2   LD3
ses_17               0.42 -1.19  1.06
acceptance_concern   0.06  0.58 -0.35
conceal_motiv       -0.23  0.05  0.32
id_uncertain         0.46 -0.29  0.15
internalized_stigma  0.18  0.09  0.08
difficult_process    0.00 -0.09  0.01
id_superior         -0.19 -0.52 -0.13
id_aff              -0.13  0.39  0.35
id_central          -0.36 -0.68 -0.28
ses_score            0.11  0.76 -1.95
cts_score            0.35 -0.23  0.78
ia_score            -0.14  0.01 -0.15
[1] "Normalized Coefficients"
                      LD1   LD2   LD3
ses_17               0.13 -0.37  0.33
acceptance_concern   0.02  0.18 -0.11
conceal_motiv       -0.07  0.02  0.10
id_uncertain         0.14 -0.09  0.05
internalized_stigma  0.06  0.03  0.02
difficult_process    0.00 -0.03  0.00
id_superior         -0.06 -0.16 -0.04
id_aff              -0.04  0.12  0.11
id_central          -0.11 -0.21 -0.09
ses_score            0.03  0.24 -0.61
cts_score            0.11 -0.07  0.24
ia_score            -0.04  0.00 -0.05
```

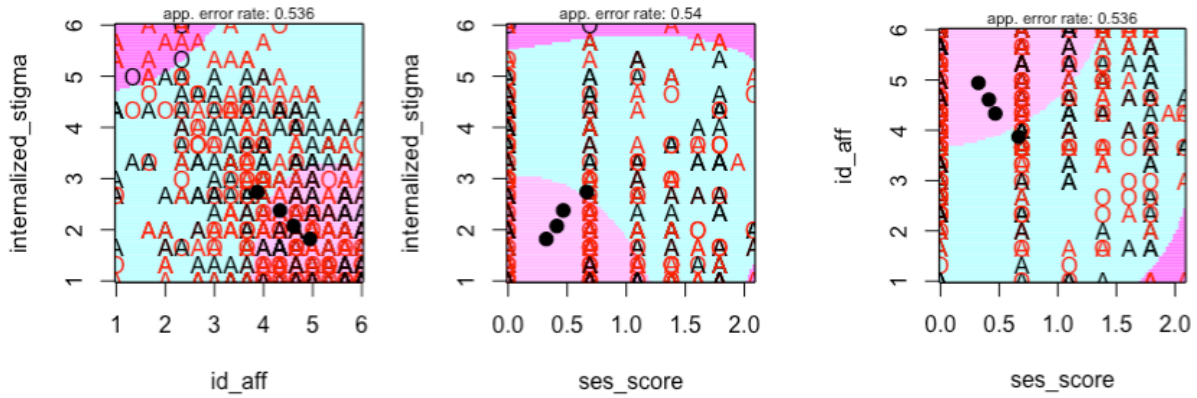
We went on to conduct both linear and quadratic DA. While the bulk of this section will be occupied with a discussion of quadratic DA, we have included both our raw/unstandardized and normalized linear DA coefficients at left. In all, it looks like sexual trauma and IPV are most central to the first discriminating function, while identity-related constructs become a bit more important in the second function (particularly acceptance concerns and identity centrality). This is especially interesting as these are two variables that have not featured prominently elsewhere in this project, compared to constructs like internalized stigma or affirmation.

Quadratic DA, however, yielded vastly different results. We utilized a stepwise DA process to determine which variables to retain in our quadratic model. Ultimately, we found that only one variable — internalized stigma — should stay. Unlike linear DA, it did not appear that variables like sexual trauma or acceptance concerns mattered significantly. While it would have been possible to continue with quadratic DA using only one variable, we decided, independently of our stepwise process, to make the model a bit more complex. Based on past analyses, we

chose to add identity affirmation and sexual trauma to the model and, upon testing for the significance of our discriminating functions, got the result below. Based on this output, we concluded that only one QDA function was significant.

Test of Function(s) <chr>	Wilks Lambda <dbl>	Approximate F <dbl>	p-value <dbl>
1 through 3	0.8809	10.6931	0.0000
2 through 3	0.9984	0.2931	0.8826
3	0.9999	0.0849	0.7708

Next, we wanted to assess our classification results. How well does our quadratic DA model (incorporating the two additional variables) predict group membership along sexual orientation lines? The answer, unfortunately, is not encouraging — under a "regular" evaluation, our model was correct only 46% of the time. This worsened marginally under cross-validated evaluation and decreased to 45%. In all, we felt that this may be a reflection of our failure to meet some of the assumptions and requirements of DA (as our Box's M test had indicated). Our partition plots underscore this, as our error rates are very high. While they are not all that telling, given these shortcomings, we have included our quadratic DA partition plots below:



These plots are difficult to read, and it is visually evident that our quadratic DA model was not all that effective. Despite the transformations we made to our outcome variables, our data still fell short of achieving multivariate normal distributions within each sexual orientation group; moreover, we rejected the null hypothesis that the covariance matrices across groups would be the same through our Box's M test. We also attempted to generate a more thorough or comprehensive model by adding additional variables to our quadratic DA, but this approach was still unsuccessful. In all, discriminant analysis was not very well-suited to our data, and it is not possible to draw any meaningful conclusions from this particular analysis. Mentioned only briefly, our linear DA coefficients may be the most substantive results we gained from discriminant analysis as we were able to interpret the discriminating functions and their key components more clearly.

Conclusion

With this project, we hoped to build upon the work done in our group member's senior thesis through various multivariate analyses. We asked two main research questions. First, is it possible to identify the key drivers of relational violence in the ace-aro community? Second, do significant differences in group means exist within this community along the lines of sexual orientation and/or gender identity? We had hoped that generalized linear models and factor analysis would aid in answering the first question; ultimately, however, our GLM analyses and results must be taken critically as our data did not quite meet the necessary assumptions. Additionally, factor analysis revealed that the outcome variables are not necessarily related to any identity-related constructs, such as internalized stigma or identity affirmation, through any deeper underlying factors. Rather, outcome variables are largely related to *each other,* and the same is true of our predictors. Finally, discriminant analysis fell short of providing us with

meaningful results as our dataset was ultimately a poor fit for that technique. Below, we have provided concluding comments and summaries of each of our three techniques.

MANOVA & Multivariate GLMs. We used MANOVA to examine group mean differences along the lines of sexual orientation and gender identity. We found stronger statistical evidence for differences across sexual orientation groups than gender identity. However, sexual orientation may not necessarily be a significant predictor of risk for relational harm: upon fitting a multivariate GLM, we found that continuous predictors, such as internalized stigma and identity affirmation, were significant across all four multivariate tests (and categorical predictors were not). This, of course, had to be interpreted with great caution as a chi-square residual plot demonstrated that our model assumptions may not have been properly met. Regarding the research questions, this set of analyses provided some support for the existence of differences in experiences between Ace-Aro individuals and those identifying with other spectrum identities (such as gray-asexuality or demisexuality); and while identity-related constructs may play a role in predicting risk, our issues with model assumptions make this difficult to claim.

Factor Analysis. Using a scree plot and parallel analysis, we determined that there might be 3 underlying factors within our dataset. However, upon describing these factors through an assessment of their loadings and loading plots, we found that they were ultimately not as informative towards our research question as we hoped. The outcome variables were all highly associated with each other and effectively comprised the first factor on their own; the continuous, identity-related variables, conversely, comprised the second two factors. While these results are intuitive, they fall short of revealing any deeper underlying "stories" linking identity experiences and relational violence. These results must also be interpreted very carefully, as our RMSR and

KMO calculations indicate that our analysis was imperfect and our dataset may not have been a perfect fit for FA, despite the variable transformations we made.

Discriminant Analysis. After making initial variable transformations and attempting both linear and quadratic discriminant analysis, we ultimately concluded that this technique was not very informative to our research questions. Our dataset was not a good fit for DA, as in addition to only roughly approximating multivariate normality within each group, there was also insufficient evidence to assume that covariance matrices were equal. In the end, our quadratic DA model yielded correct classifications less than 50% of the time. While DA was intended to help answer our second research question, MANOVA provided better results to that end.

Limitations & Directions for Future Study. This project contained major limitations in that the dataset used was not entirely compatible with the assumptions of several multivariate analyses. It is possible that with further and more creative variable transformations, beyond the log transforms undertaken here, better results may have been obtained. However, future studies similarly working with public health data may need to plan upfront for such limitations and design surveys strategically such that the data yielded will be more appropriate.