Discriminant Analysis

Chichang Xiong

Kelly Kinahan

COM 631

March 27, 2013

I. Model

Using the Humor and Public Opinion Data Set (Neuendorf & Skalski, 2010)

IVs:

C44 reverse coded

C17

C22

C23

C27 reverse coded

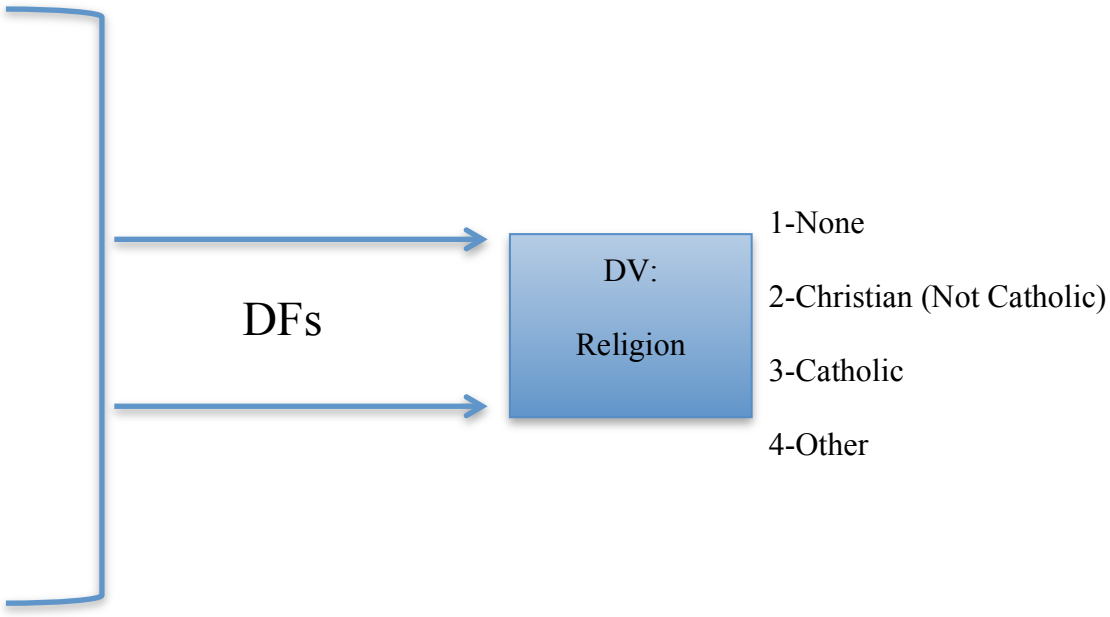
C30

C45

C41

C28

C18 reverse coded



Key:

C44 reverse coded- I like satire.

C17 - I like it when friends give each other a hard time by joking with them.

C22 - I like sarcasm.

C23 - I like humor that is naughty.

C27 reverse coded - I like humor that is delivered in a dry manner.

C30 - I like humor that puts down arrogant people.

C45 - I like when people joke around socially to have fun.

C41 - I like dark comedy.

C28 - I like humor that puts down men.

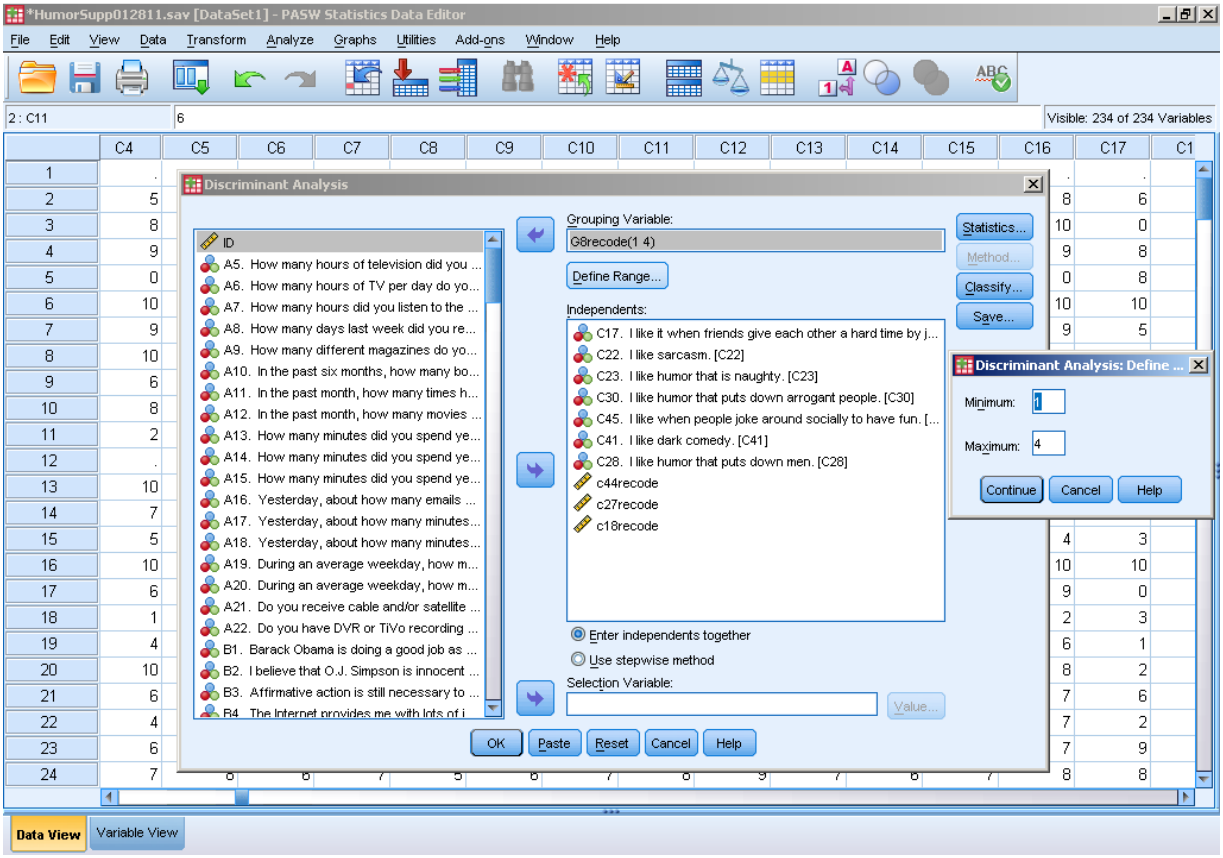
C18 reverse coded- I like humor that puts down women.

II. Running SPSS

To perform Discriminant Function Analysis:

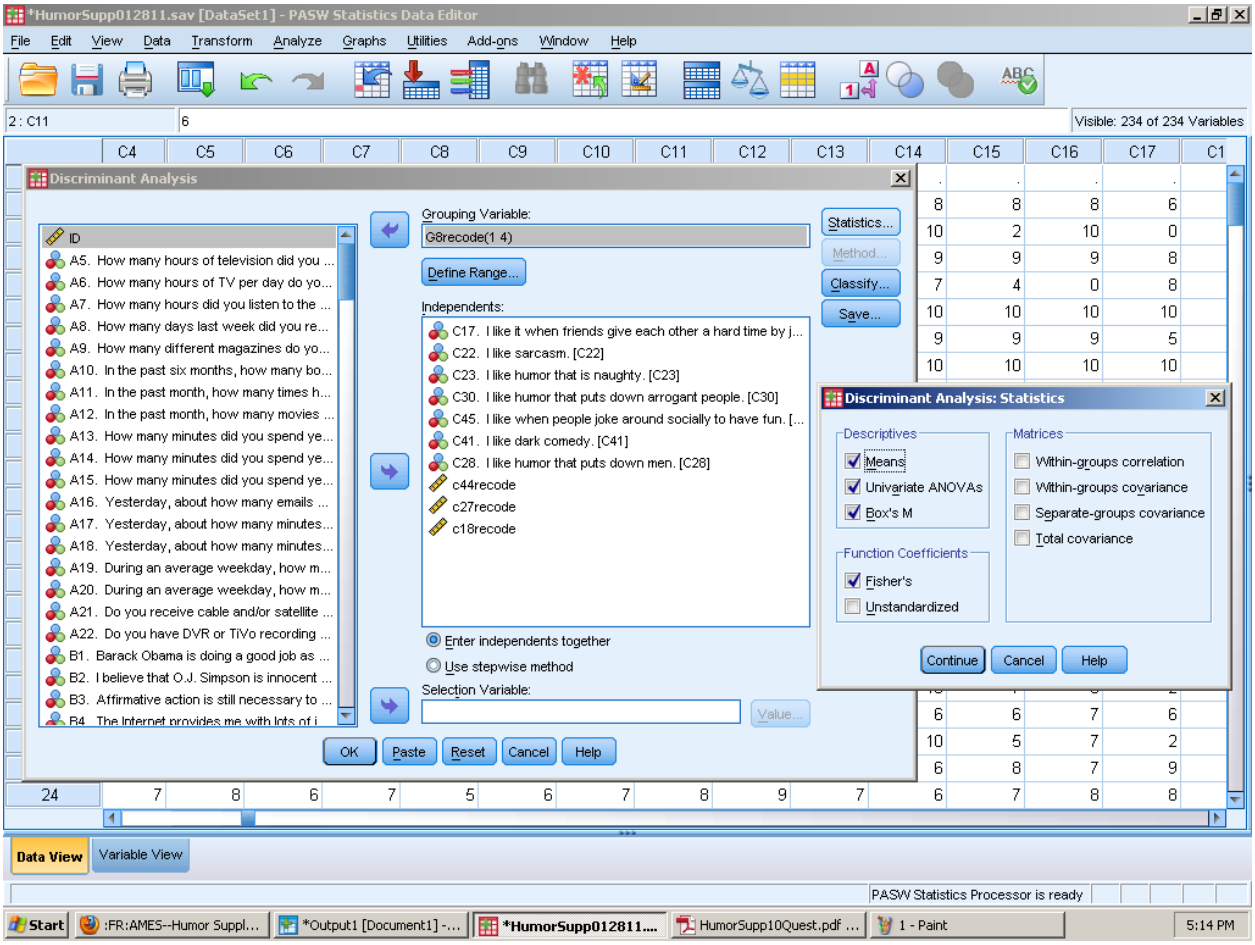
Analyze→ Classify→ Discriminant

- Pick your DV from the left column and click the arrow to bring it into the box labeled Grouping Variable.
- Click on Define Range and identify the minimum and maximum values (in this case, 1 and 4).
- Click Continue.
- Pick your IVs from the left column and click the arrow to bring them into the box labeled Independents.
- Underneath the Independents box, select Enter Independents Together (forced entry, not stepwise).



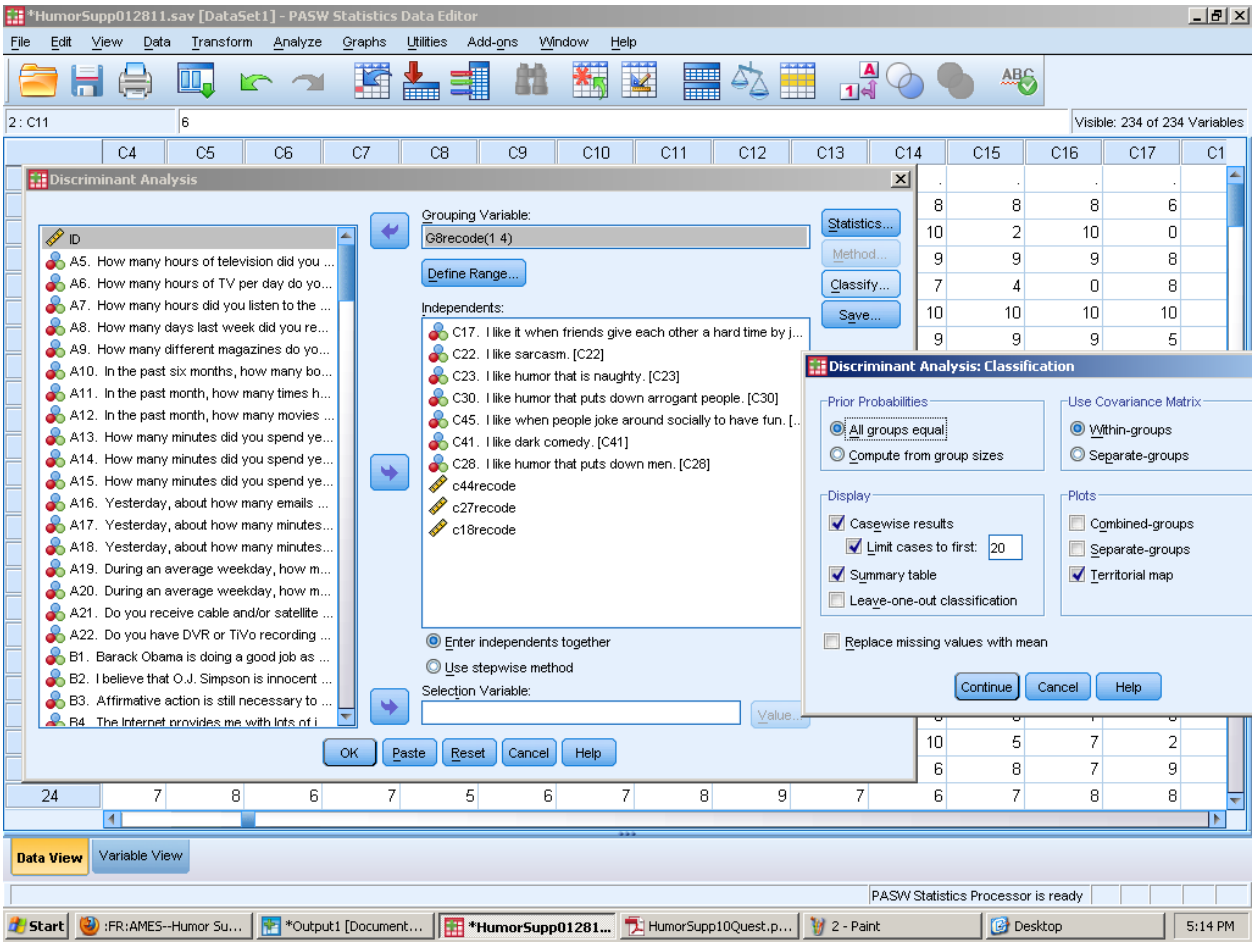
To perform Discriminant Function Analysis:

- Click on the **Statistics** button.
- In the **Discriminant Analysis: Statistics** window, select **Means**, **Univariate ANOVAs**, and **Box's M**.
- Under **Functions Coefficients** check **Fisher's**.
- Click **Continue**.



To perform Discriminant Function Analysis:

- Click on **Classify**.
- Under **Prior Probabilities**, choose **All Groups Equal**.
- Under **Display**, select **Casewise Results**, **Limit Cases to First 20**, and **Summary Table**.
- Under **Use Covariance Matrix**, choose **Within-Groups**.
- Under **Plots**, select **Territorial Map**.
- Click **Continue** and **OK** to run the **Discriminant Analysis** output.



III. SPSS Output

```
DISCRIMINANT
  /GROUPS=G8recode(1 4)
  /VARIABLES=C17 C22 C23 C30 C45 C41 C28 c44recode c27recode
c18recode
  /ANALYSIS ALL
  /PRIORS EQUAL
  /STATISTICS=MEAN STDDEV UNIVF BOXM COEFF TABLE
  /PLOT=MAP
  /PLOT=CASES(20)
  /CLASSIFY=NONMISSING POOLED.
```

Discriminant

```
[DataSet1] C:\DOCUME~1\2576279\LOCALS~1\Temp\HumorSupp012811.sav
```

Analysis Case Processing Summary

Unweighted Cases		N	Percent
Valid		225	78.1
Excluded	Missing or out-of-range group codes	32	11.1
	At least one missing discriminating variable	11	3.8
	Both missing or out-of-range group codes and at least one missing discriminating variable	20	6.9
	Total	63	21.9
Total		288	100.0

Group Statistics

G8recode (religious affiliation)		Mean	Std. Deviation	Valid N (listwise)	
				Unweighted	Weighted
1.00	C17. I like it when friends give each other a hard time by joking with them.	6.6867	2.52754	83	83.000
	C22. I like sarcasm.	8.5060	2.17734	83	83.000
	C23. I like humor that is naughty.	7.0843	2.62820	83	83.000
	C30. I like humor that puts down arrogant people.	7.3855	2.37790	83	83.000
	C45. I like when people joke around socially to have fun.	8.3614	1.76399	83	83.000
	C41. I like dark comedy.	6.4699	2.93148	83	83.000
	C28. I like humor that puts down men.	4.5904	3.06449	83	83.000
	c44 reverse coded. Like satire.	6.4217	3.23140	83	83.000
	c27 reverse coded. Like dry humor.	6.5181	3.03380	83	83.000
	c18 reverse coded. Like humor putting down women.	4.4217	3.36815	83	83.000
2.00	C17. I like it when friends give each other a hard time by joking with them.	6.0154	3.15977	65	65.000
	C22. I like sarcasm.	7.9692	2.58583	65	65.000
	C23. I like humor that is naughty.	6.4462	2.92099	65	65.000
	C30. I like humor that puts down arrogant people.	7.1231	2.36846	65	65.000
	C45. I like when people joke around socially to have fun.	7.6615	2.34705	65	65.000
	C41. I like dark comedy.	4.4462	3.14260	65	65.000

	C28. I like humor that puts down men.	4.8462	3.23666	65	65.000
	c44 reverse coded. Like satire.	7.1846	2.51180	65	65.000
	c27 reverse coded. Like dry humor.	5.2154	3.54212	65	65.000
	c18 reverse coded. Like humor putting down women.	4.8769	3.18938	65	65.000
3.00	C17. I like it when friends give each other a hard time by joking with them.	6.6308	2.42751	65	65.000
	C22. I like sarcasm.	7.3538	2.49017	65	65.000
	C23. I like humor that is naughty.	6.7231	2.34859	65	65.000
	C30. I like humor that puts down arrogant people.	6.8769	2.72436	65	65.000
	C45. I like when people joke around socially to have fun.	7.5846	2.31093	65	65.000
	C41. I like dark comedy.	4.9846	2.67215	65	65.000
	C28. I like humor that puts down men.	4.2769	2.77549	65	65.000
	c44 reverse coded. Like satire.	6.8000	2.61127	65	65.000
	c27 reverse coded. Like dry humor.	5.5538	2.85608	65	65.000
	c18 reverse coded. Like humor putting down women.	4.2000	2.55074	65	65.000
4.00	C17. I like it when friends give each other a hard time by joking with them.	7.2500	3.04884	12	12.000
	C22. I like sarcasm.	9.0000	1.80907	12	12.000
	C23. I like humor that is naughty.	6.2500	3.76889	12	12.000

	C30. I like humor that puts down arrogant people.	8.4167	1.72986	12	12.000
	C45. I like when people joke around socially to have fun.	7.8333	2.08167	12	12.000
	C41. I like dark comedy.	5.9167	3.67939	12	12.000
	C28. I like humor that puts down men.	3.9167	2.87492	12	12.000
	c44 reverse coded. Like satire.	7.1667	2.85509	12	12.000
	c27 reverse coded. Like dry humor.	5.1667	3.27062	12	12.000
	c18 reverse coded. Like humor putting down women.	6.2500	3.98006	12	12.000
Total	C17. I like it when friends give each other a hard time by joking with them.	6.5067	2.72901	225	225.000
	C22. I like sarcasm.	8.0444	2.41790	225	225.000
	C23. I like humor that is naughty.	6.7511	2.70587	225	225.000
	C30. I like humor that puts down arrogant people.	7.2178	2.46251	225	225.000
	C45. I like when people joke around socially to have fun.	7.9067	2.13918	225	225.000
	C41. I like dark comedy.	5.4267	3.06990	225	225.000
	C28. I like humor that puts down men.	4.5378	3.01775	225	225.000
	c44 reverse coded. Like satire.	6.7911	2.84195	225	225.000
	c27 reverse coded. Like dry humor.	5.7911	3.18421	225	225.000
	c18 reverse coded. Like humor putting down women.	4.5867	3.14994	225	225.000

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
C17. I like it when friends give each other a hard time by joking with them.	.984	1.167	3	221	.323
C22. I like sarcasm.	.954	3.538	3	221	.016
C23. I like humor that is naughty.	.989	.832	3	221	.477
C30. I like humor that puts down arrogant people.	.980	1.535	3	221	.206
C45. I like when people joke around socially to have fun.	.973	2.060	3	221	.106
C41. I like dark comedy.	.920	6.386	3	221	.000
C28. I like humor that puts down men.	.992	.563	3	221	.640
c44 reverse coded (satire)	.987	.952	3	221	.416
c27 reverse coded (dry humor)	.968	2.472	3	221	.063
c18 reverse coded (put down women)	.977	1.718	3	221	.164

Analysis 1

Box's Test of Equality of Covariance Matrices

Log Determinants

G8recode	Rank	Log Determinant
1.00	10	17.120
2.00	10	18.762
3.00	10	16.823
4.00	10	11.997
Pooled within-groups	10	18.492

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

Test Results

Box's M		273.426
F	Approx.	1.358
	df1	165
	df2	5432.824
	Sig.	.002

Tests null hypothesis of equal population covariance matrices.

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.170[a]	56.3	56.3	.382
2	.088[a]	29.1	85.5	.285
3	.044[a]	14.5	100.0	.205

a. First 3 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 3	.752	61.779	30	.001
2 through 3	.880	27.642	18	.068
3	.958	9.320	8	.316

Standardized Canonical Discriminant Function Coefficients

	Function		
	1	2	3
C17. I like it when friends give each other a hard time by joking with them.	.070	-.195	-.940
C22. I like sarcasm.	.172	.717	.451
C23. I like humor that is naughty.	-.004	-.606	.195
C30. I like humor that puts down arrogant people.	.022	.439	-.258
C45. I like when people joke around socially to have fun.	.332	-.039	.189
C41. I like dark comedy.	.649	.073	-.257
C28. I like humor that puts down men.	-.230	-.091	.735
c44 reverse coded. Like satire.	-.447	-.007	-.102
c27 reverse coded. Like dry humor.	.458	-.244	.307
c18 reverse coded. Like humor putting down women.	-.220	.621	-.061

Structure Matrix

	Function		
	1	2	3
C41. I like dark comedy.	.705*	.126	-.110

c27 reverse coded(dry humor)	.430*	-.110	.149
C45. I like when people joke around socially to have fun.	.381*	.101	.227
c44 reverse coded (satire)	-.258*	.133	.006
C23. I like humor that is naughty.	.233*	-.146	.070
C22. I like sarcasm.	.334	.551*	.230
c18 reverse coded (put down women)	-.097	.492*	-.093
C30. I like humor that puts down arrogant people.	.148	.429*	-.139
C17. I like it when friends give each other a hard time by joking with them.	.201	.034	-.449*
C28. I like humor that puts down men.	-.032	.018	.411*

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions

Variables ordered by absolute size of correlation within function.

*. Largest absolute correlation between each variable and any discriminant function

Functions at Group Centroids

G8recode (religion)	Function		
	1	2	3
1.00 (none)	.510	.009	.082
2.00 (other Christians)	-.461	.136	.205
3.00 (Catholics)	-.186	-.327	-.209
4.00 (other)	-.024	.971	-.544

Unstandardized canonical discriminant functions evaluated at group means

Classification Statistics

Classification Processing Summary

Processed		288
Excluded	Missing or out-of-range group codes	0
	At least one missing discriminating variable	31
Used in Output		257

Prior Probabilities for Groups

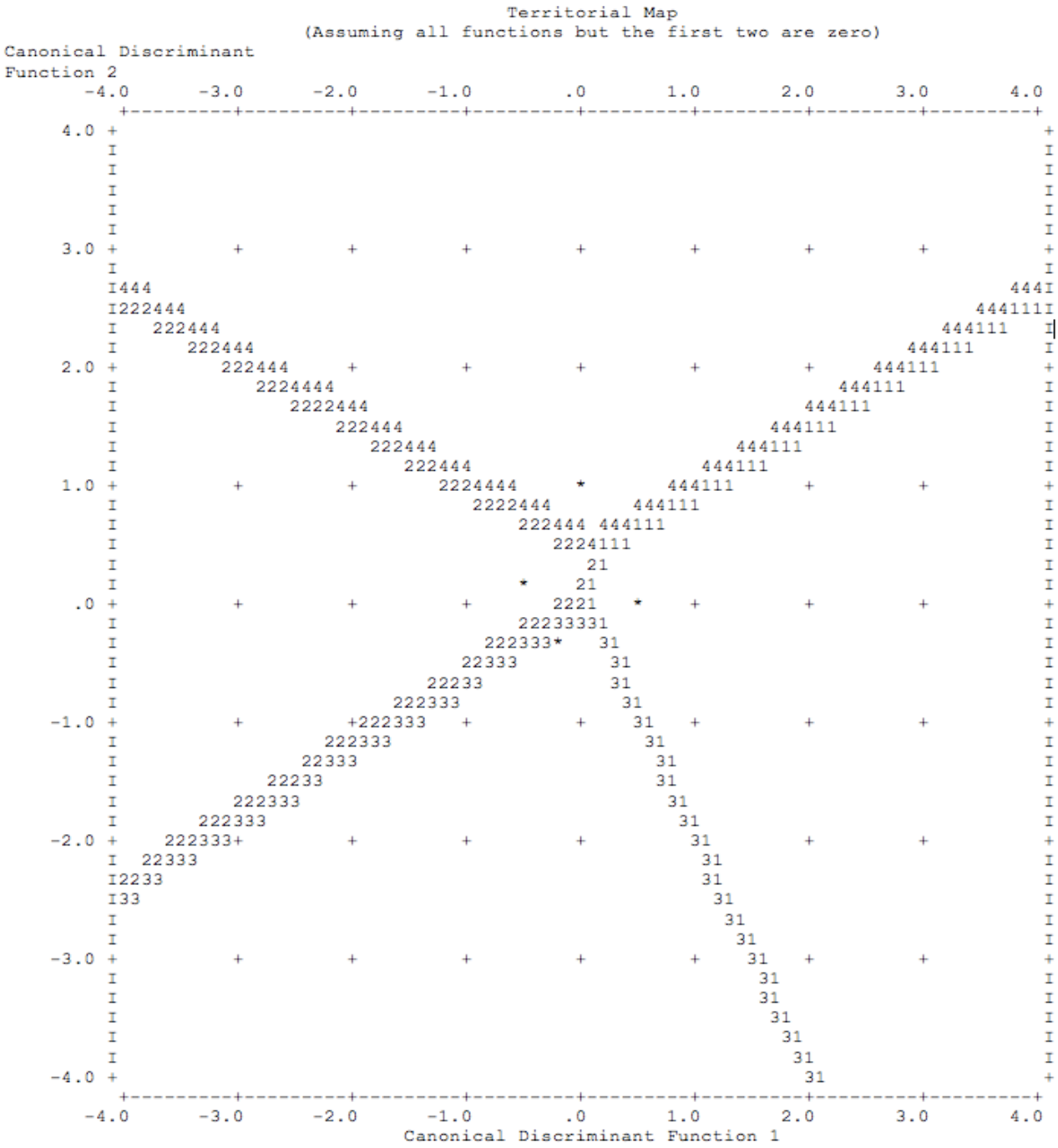
G8recode		Cases Used in Analysis	
	Prior	Unweighted	Weighted
1.00	.250	83	83.000
2.00	.250	65	65.000
3.00	.250	65	65.000
4.00	.250	12	12.000
Total	1.000	225	225.000

Classification Function Coefficients

	G8recode			
	1.00	2.00	3.00	4.00
C17. I like it when friends give each other a hard time by joking with them.	.190	.113	.297	.323
C22. I like sarcasm.	.509	.501	.302	.642
C23. I like humor that is naughty.	-.085	-.103	-.029	-.344
C30. I like humor that puts down arrogant people.	.669	.670	.633	.902
C45. I like when people joke around socially to have fun.	1.256	1.113	1.128	1.100
C41. I like dark comedy.	.299	.078	.163	.259

C28. I like humor that puts down men.	.080	.180	.072	-.061
c44 reverse coded. Like satire.	.382	.531	.503	.486
c27 reverse coded. Like dry humor.	.630	.492	.527	.417
c18 reverse coded. Like humor putting down women.	.202	.293	.190	.442
(Constant)	-16.488	-14.551	-13.970	-17.326

Fisher's linear discriminant functions



Symbols used in territorial map

Symbol	Group	Label
1	1	
2	2	
3	3	
4	4	

* Indicates a group centroid

Casewise Statistics

	Case Number	Actual Group	Highest Group					Second Highest Group			Discriminant Scores		
			Predicted Group	P(D>d | G=g)		P(G=g | D=d)	Squared Mahalanobis Distance to Centroid	Group	P(G=g | D=d)	Squared Mahalanobis Distance to Centroid	Function 1	Function 2	Function 3
				p	df								
Original	2	3	1**	.624	3	.343	1.760	2	.330	1.839	.126	-.523	1.235
	3	2	2	.064	3	.696	7.252	3	.174	10.024	-1.652	-.218	2.595
	4	ungrouped	1	.967	3	.392	.263	3	.303	.775	.678	-.444	-.086
	5	4	4	.308	3	.698	3.599	2	.193	6.166	-.863	2.579	.011
	6	3	3	.011	3	.632	11.193	1	.297	12.706	1.153	-2.941	-1.812
	7	1	1	.051	3	.710	7.786	2	.155	10.828	1.843	-.304	2.513
	8	1	1	.815	3	.428	.945	3	.328	1.478	.948	-.742	-.351
	9	2	2	.918	3	.375	.502	3	.273	1.137	-.309	-.152	.835
	10	ungrouped	3	.707	3	.478	1.393	2	.294	2.366	-.645	-1.383	.049
	11	3	3	.396	3	.536	2.973	2	.301	4.125	-1.486	-1.220	-.906
	12	2	2	.857	3	.481	.769	3	.262	1.987	-1.060	.030	.837
	13	2	3**	.963	3	.357	.286	2	.287	.725	-.574	-.218	-.561
	14	2	1**	.971	3	.384	.237	4	.217	1.375	.734	.432	.171
	15	ungrouped	2	.948	3	.313	.361	1	.282	.571	.001	.375	.505
	16	ungrouped	3	.056	3	.593	7.566	2	.192	9.824	-1.583	-1.085	-2.454
	17	2	2	.146	3	.507	5.375	3	.229	6.964	-2.641	.617	-.419
	18	4	4	.808	3	.674	.972	2	.151	3.969	-.412	1.822	-.854
	20	2	1**	.754	3	.362	1.197	2	.294	1.612	.266	-.546	.992
	21	3	4**	.949	3	.311	.355	3	.253	.766	-.142	.405	-.687
	22	1	4**	.708	3	.715	1.391	2	.133	4.753	-.249	2.122	-.669

**. Misclassified case

Classification Results[a]

		G8recode	Predicted Group Membership				
			1.00	2.00	3.00	4.00	Total
Original	Count	1.00	38	11	18	16	83
		2.00	18	26	12	9	65
		3.00	18	8	26	13	65
		4.00	2	2	2	6	12
		Ungrouped cases	10	6	13	3	32
	%	1.00	45.8	13.3	21.7	19.3	100.0
		2.00	27.7	40.0	18.5	13.8	100.0
		3.00	27.7	12.3	40.0	20.0	100.0
		4.00	16.7	16.7	16.7	50.0	100.0
		Ungrouped cases	31.3	18.8	40.6	9.4	100.0

a. 42.7% of original grouped cases correctly classified.

IV. Tabling Results:

Table 1

			Loadings	
IVs	DF1-"Edgy" Standardized Coefficients	DF2-"Cynical" Standardized Coefficients	DF1-"Edgy" Correlation	DF2-"Cynical" Correlation
C17-like joking with friends	.070	-.195	.201	.034
C22-like sarcasm	.172	.717	.334	.551*
C23-like naughty humor	-.004	-.606	.233*	-.146
C30-like humor putting down arrogant people	.022	.439	.148	.429*
C45-like joking socially	.332	-.039	.381*	.101
C41-like dark comedy	.649	.073	.705*	.126
C28-like humor putting down men	-.230	-.091	-.032	.018
C44 reverse code-like satire	-.447	-.007	-.258*	.133
C27 reverse code-like dry humor	.458	-.244	.430*	-.110
C18 reverse code-like humor putting down women	-.220	.621	-.097	.492*

*Indicates largest correlation between each variable and any discriminant function.

Table 2
Mean scores on discriminant function for 4DV groups (centroids)

Religion	DF1-"Edgy":	DF2 –"Cynical":
1-None	.510	.009
2-Christian (not Catholic)	-.461	.136
3-Catholic	-.186	-.327
4-Other	-.024	.971
Wilks' Lambda	.752	.880
Chi Square	61.779	27.642
Significance	.001	.068
Eigen value	.170	.088
Canonical Correlation	.382	.285

Table 3
Classification Matrix results for 4 group discriminant analysis

Actual Group		Predicted Group			
Group	Actual Group Size	1-None	2-Christian (not Catholic)	3-Catholic	4-Other
1-None	83	38	11	18	16
2-Christian (not Catholic)	65	18	26	12	9
3-Catholic	65	18	8	26	13
4-Other	12	2	2	2	6
Total	225				

42.7% of original grouped cases correctly classified.

Press' Q (tests whether the classification analysis improves prediction to groups significantly):

$$\frac{[N-(nK)]2}{N(K-1)}$$

N=225

n=96

K=4

$$\frac{[225-(96*4)]2}{225*(4-1)}$$

$$=\frac{(225-384)2}{675}$$

$$=\frac{25281}{675}$$

=37.45

Critical Value when df=1 on chi square table is 6.63, but our value 37.45 exceeds the critical value, indicating SIGNIFICANT at $p<0.001$.

V. Write-up of Results

A discriminant function analysis was applied to assess how well an individual's religion could be predicted from 10 items from the Humor and Public Opinion dataset. These ten discriminating independent variables include: do not like satire, I like it when friends give each other a hard time by joking with them, I like sarcasm, I like humor that is naughty, I do not like humor that is delivered in a dry manner, I like humor that puts down arrogant people, I like when people joke around socially to have fun, I like dark comedy, I like humor that puts down men, and I do not like humor that puts down women. The "not like" variables were reverse coded. The dependent variable is religion, and was recoded from seven original options that received answers to four groups that reflected a better distribution of the data. These groups included: none, Christian (not Catholic) which was a combination of Protestant and Other Christian, Catholic, and Other which was a combination of Muslim, Jewish, and Buddhist.

This analysis produced three discriminant functions, one that was significant ($p<.001$) and one that was near significant ($p<.10$). The first discriminant function was labeled "Edgy" because the variables that loaded highly on this function were thought to be edgier types of humor (dark [.705], dry [.430], naughty [.233], socially joking with friends [.381], and disliking satire [-.258]). The Wilks' Lambda, examines how much the groups differ on the set of independent variables, is .752 for the first discriminant function. The second discriminant function was labeled "Cynical" because the variables that loaded highly on this function included sarcasm (.551), putting down women (.492), and putting down arrogant people (.429). The Wilks' Lambda of the second discriminant function (.880) is greater than that of the first, reflective of its weaker discriminating ability.

Table 2 reflects the mean scores for each of the four dependent variable groups on the

two discriminant functions. The group centroids show a pattern that suggests those with no religious affiliation (Group 1) like "Edgy" humor, while Christians (not Catholics) tend to not like this type of humor. Group 1 (no religious affiliation) has positive and highest means on the "Edgy" discriminant function (dark, dry, joke socially, don't like satire, naughty) and Group 2 (other Christians not Catholics) has negative and largest absolute value of means on this discriminant function. Furthermore, Catholics tend to dislike "Cynical" humor, while Other religions (generally minorities) like this type of humor. Group 3 (Catholics) has negative and largest absolute value of means on the "Cynical" (sarcasm, put down women, put down arrogant people) and Group 4 (other, aka minorities) has positive and highest means on the "Cynical" discriminant function.

However, from this analysis, while we can assess that Group 1 (no religious affiliation) and Group 4 (other religions) have highest means on discriminant functions one ("Edgy") and two ("Cynical") respectively, we cannot say that Group 1 and Group 2 have SIGNIFICANT higher means than other groups on DF1 and DF2. To tell whether it's significant or not, we could further conduct a post-hoc test (in ANOVA).

As shown in Table 3, of all the cases in total 42.7% could be correctly classified into the 4 religion groups of the DV by our discriminant analysis. The Press'Q was calculated at 37.45, which is bigger than the critical value of 6.63 ($df=1$, $p<.001$), indicating that using the IVs that we chose to predict religion groups are significantly more useful than by chance.