

Canonical Correlation

Christopher Lohr
COM 631
April 15, 2014

PART I: MODEL

Canonical correlation is typically treated as an exploratory process whereby two sets of variables are analyzed, with patterns of linear relationships between the two sets discovered via newly created variates.

It can be related to Factor Analysis in that it unpacks overarching “themes” among the variables (within a given set) – themes that can themselves be named.

For this analysis, two sets of variables were used from the Humor & Public Opinion dataset (Neuendorf et al.) These are:

SET 1 – “Internalized Values”	SET 2 – “Externalized Values”
Inner Harmony (F2)	Social Power (F3)
A Spiritual Life (F6)	Social Order (F8)
Sense of Belonging (F7)	Wealth (F12)
Meaning in Life (F10)	Respect for Tradition (F18)
Self-Respect (F14)	Authority (F27)
Self-Discipline (F20)	Honoring of Parents & Elders (F40)
Wisdom (F26)	

These variables were chosen from among the Schwartz Values Scales questions (F questions in the dataset) but do not correspond precisely with any of the defined scales, i.e. benevolence, hedonism, power, etc. Instead these variables were hypothesized as belong to either:

- internalized values (SET 1) or;
- externalized values (SET 2)

In each case the item was measured from -1 to 7 with 0 indicating not important as a guiding principle in the respondent’s life, 7 indicating of supreme importance, and -1 indicating opposed to the respondent’s values.

Schwartz Values Scales (1992)

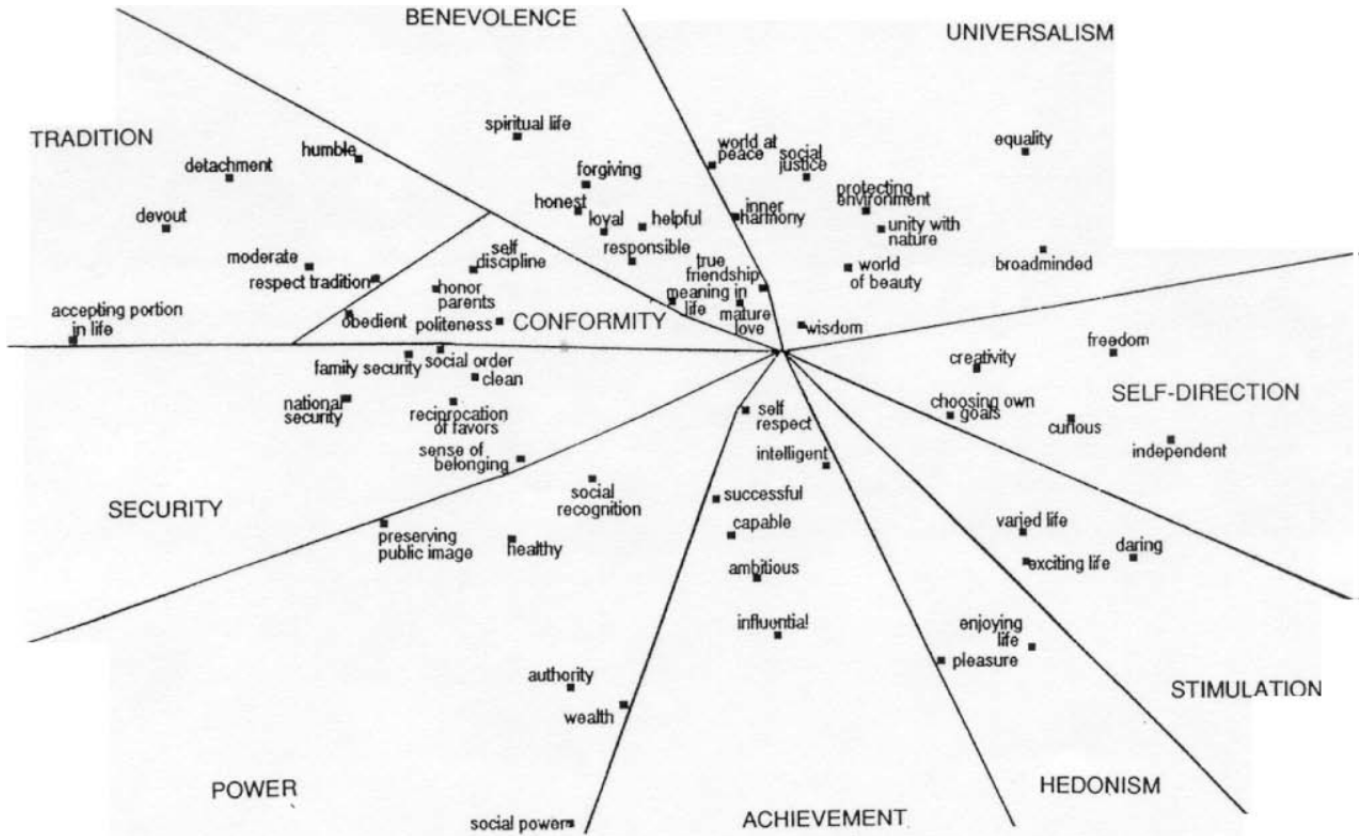
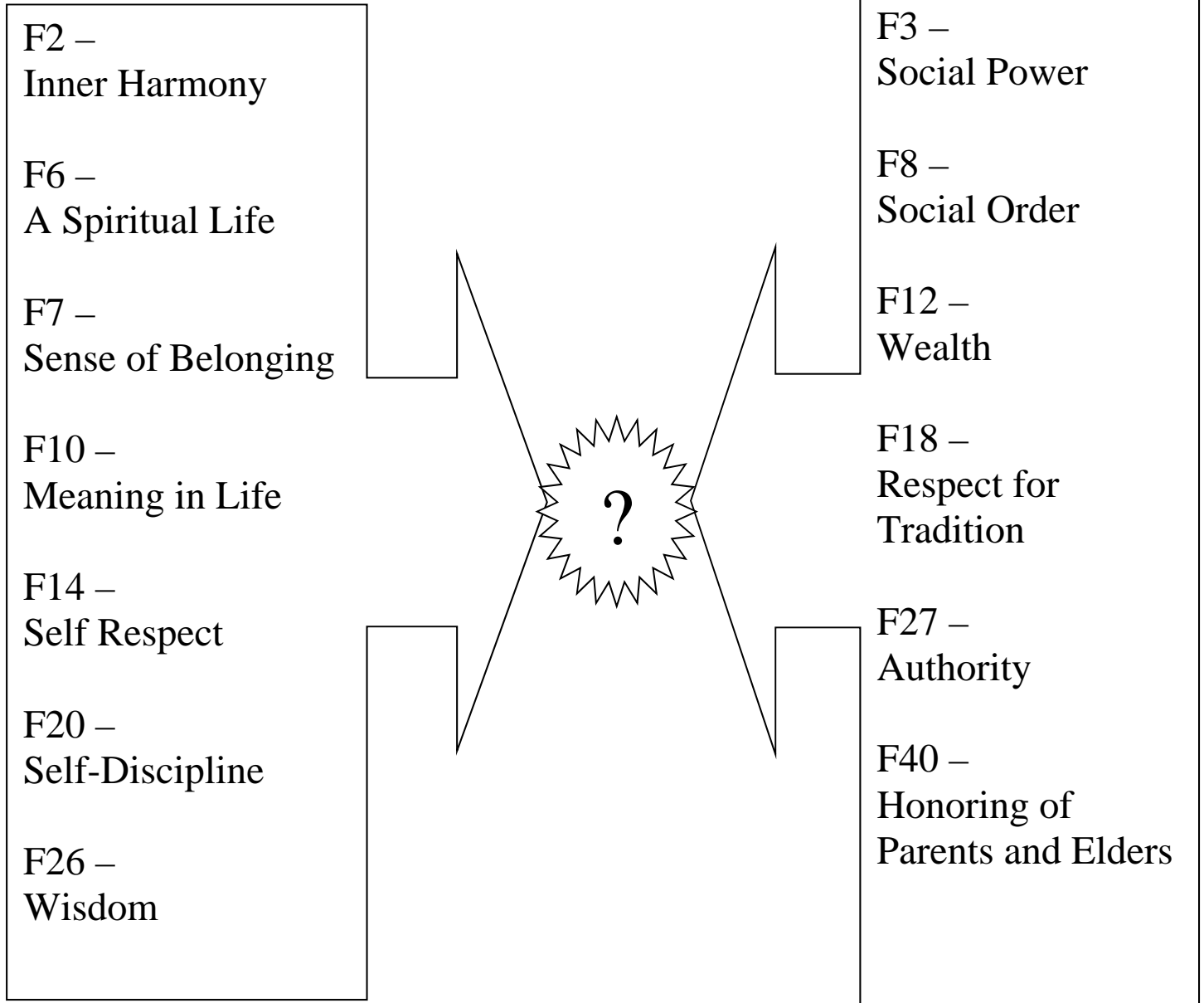


Fig. 4. Teacher samples. Individual-level value structure averaged across 15 countries: Two-dimensional smallest space analysis.

SET 1

SET 2

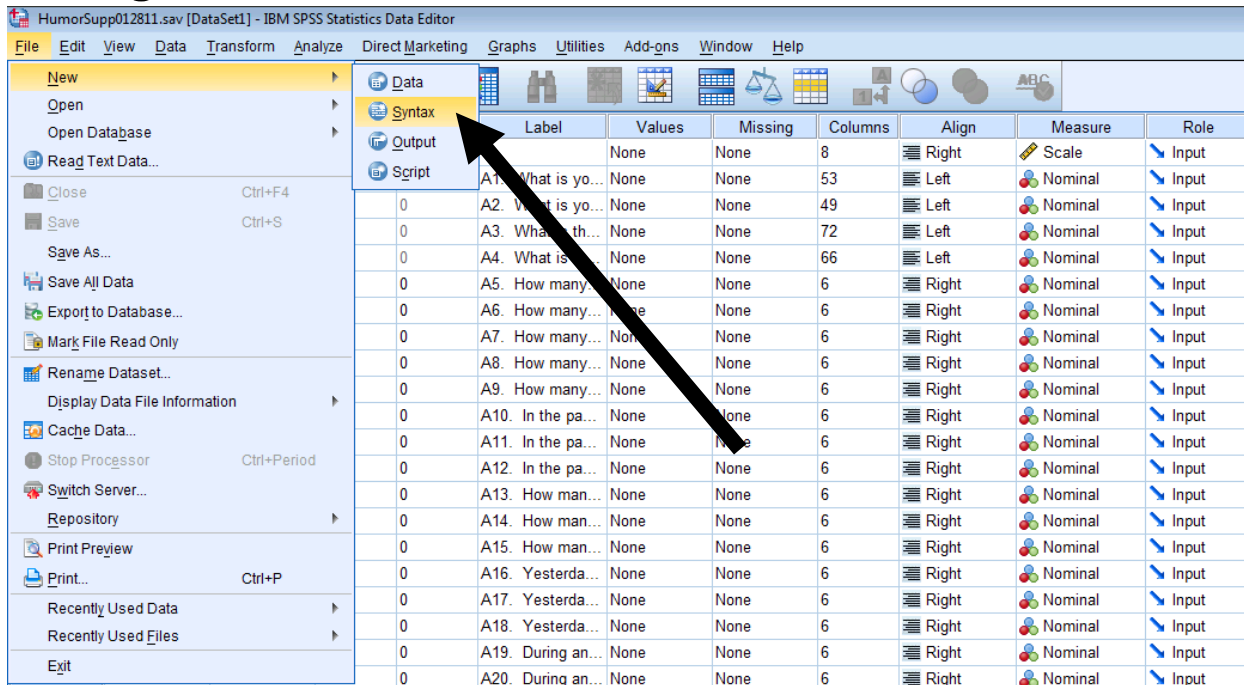


PART II: RUNNING SPSS

Unfortunately, the procedure for running a canonical correlation is not as straightforward as in some other analyses. The process must be done through SYNTAX. Let's take a look how this is done!

First, open your dataset – in this case the humor dataset.

Then go to FILE – NEW – SYNTAX

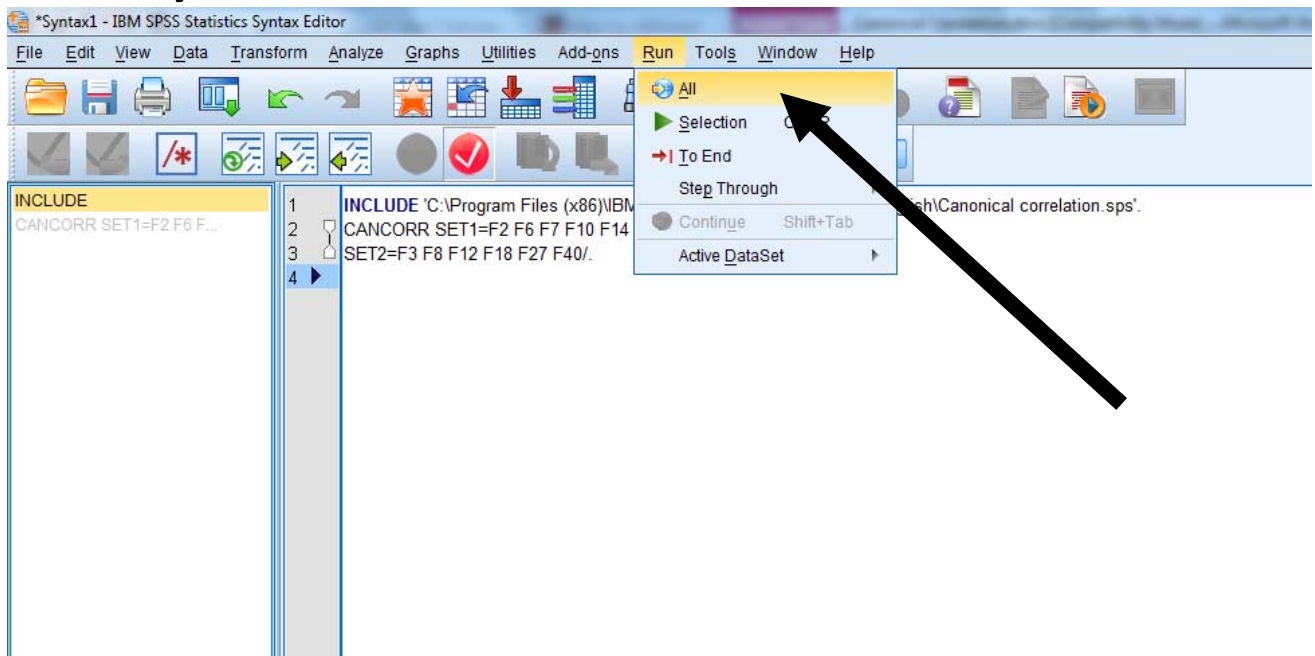


Now with the SYNTAX box open, plug in the necessary command. For Canonical Correlation this looks like:

```
INCLUDE 'C:\Program Files  
(x86)\IBM\SPSS\Statistics\20\Samples\English\Canonical  
correlation.sps'.  
CANCORR SET1=F2 F6 F7 F10 F14 F20 F26 /  
SET2=F3 F8 F12 F18 F27 F40/.
```

NOTE: Highlighted portion will vary from machine to machine – check where your canonical correlation .sps file is located to insert this. Use only one quote mark on each end.

In the syntax editor click RUN – ALL



Now hopefully it ran – the screen will usually flash and change screens away from the Syntax editor – that is fine.

You will also likely see a whole lot of error messages in your output – this is probably OK as long as the canonical correlation made it through (more on that in a moment)

The error messages will look something like this:

```
Error # 34 in column 24.  Text: canonical_correlation_size_.sav
SPSS Statistics cannot access a file with the given file
specification.  The file specification is either syntactically
invalid, specifies an invalid drive, specifies a protected
directory, specifies a protected file, or specifies a non-
sharable file.
Execution of this command stops.
```

As long as the MATRIX, which includes the canonical correlation and redundancy analyses, ran properly these errors are likely OK.

Additionally, SPSS will run a CREATE command that attempts to append the new canonical correlation variates to the existing datasets.

It is OK If this does not work (also OK if it does).

PART III: SPSS OUTPUT

GET

```
FILE='C:\Users\Asus\Downloads\HumorSupp012811.sav'.  
DATASET NAME DataSet1 WINDOW=FRONT.
```

```
INCLUDE 'C:\Program Files  
(x86)\IBM\SPSS\Statistics\20\Samples\English\Canonical  
correlation.sps'.  
5959 0 * Canonical correlation.sps. This version allows long  
variable names and uses datasets.Canonical correlation.sps.  
5960 0  
5962 0 preserve.  
5964 0 set printback=off.  
6668 0 RESTORE.  
6669 0  
6671 0 * End of INSERT and INCLUDE nesting level 01.  
CANCORR SET1=F2 F6 F7 F10 F14 F20 F26 /  
SET2=F3 F8 F12 F18 F27 F40/.
```


Canonical Correlation

Matrix

Run MATRIX procedure:

Correlations for Set-1

	F2	F6	F7	F10	F14	F20	F26
F2	1.0000	.3316	.3383	.4074	.5448	.3849	.4623
F6	.3316	1.0000	.2537	.3802	.3315	.3703	.3013
F7	.3383	.2537	1.0000	.4074	.5188	.2247	.3950
F10	.4074	.3802	.4074	1.0000	.5720	.3652	.5003
F14	.5448	.3315	.5188	.5720	1.0000	.4634	.5637
F20	.3849	.3703	.2247	.3652	.4634	1.0000	.4752
F26	.4623	.3013	.3950	.5003	.5637	.4752	1.0000

Correlations for Set-2

	F3	F8	F12	F18	F27	F40
F3	1.0000	.1646	.3712	.0900	.4671	-.0129
F8	.1646	1.0000	.4194	.3322	.4778	.3241
F12	.3712	.4194	1.0000	.1717	.5366	.1361
F18	.0900	.3322	.1717	1.0000	.3156	.5173
F27	.4671	.4778	.5366	.3156	1.0000	.3152
F40	-.0129	.3241	.1361	.5173	.3152	1.0000

Correlations Between Set-1 and Set-2

	F3	F8	F12	F18	F27	F40
F2	-.0755	.2945	.0573	.1902	.1594	.3800
F6	.0574	.3186	.0512	.3168	.1946	.3741
F7	.0618	.5863	.2352	.2580	.2710	.3346
F10	.0370	.3991	.1644	.2603	.2652	.3493
F14	-.0401	.4489	.2195	.2423	.2650	.4368
F20	.0408	.3729	.1543	.3689	.3318	.4049
F26	-.1179	.4110	.0928	.3296	.2652	.4738

Note: These are straightforward correlations between variables within and between the sets. The canonical correlations are next.

Canonical Correlation

Canonical Correlations

1	.740
2	.333
3	.268
4	.172
5	.126
6	.044

Test that remaining correlations are zero:

	Wilk's	Chi-SQ	DF	Sig.
1	.356	209.639	42.000	.000
2	.787	48.711	30.000	.017
3	.885	24.864	20.000	.207
4	.953	9.770	12.000	.636
5	.982	3.656	6.000	.723
6	.998	.392	2.000	.822

Note: Only the first two canonical correlations are significant (p < .05). These are highlighted for reference throughout

Canonical Correlation

Standardized Canonical Coefficients for Set-1

	1	2	3	4	5	6
F2	.014	.350	.316	.543	.117	-.867
F6	.213	.163	-.457	.717	.451	.517
F7	.481	-.905	-.036	.270	-.454	-.002
F10	.054	-.239	-.262	-.241	.127	-.798
F14	.076	.028	.900	-.482	.827	.711
F20	.278	-.013	-.700	-.722	.083	-.175
F26	.293	.737	.120	.077	-.938	.375

Raw Canonical Coefficients for Set-1

	1	2	3	4	5	6
F2	.012	.299	.270	.463	.100	-.741
F6	.091	.069	-.195	.306	.192	.220
F7	.309	-.582	-.023	.174	-.292	-.002
F10	.044	-.195	-.214	-.197	.103	-.651
F14	.071	.026	.837	-.449	.770	.662
F20	.191	-.009	-.481	-.496	.057	-.120
F26	.246	.620	.101	.065	-.790	.315

Standardized Canonical Coefficients for Set-2

	1	2	3	4	5	6
F3	-.115	-.496	-.534	.505	.718	-.234
F8	.676	-.704	.203	.505	-.458	-.107
F12	-.082	-.316	.521	-.662	.500	.683
F18	.112	.100	-.848	-.135	-.251	.788
F27	.112	.461	-.294	-.839	-.439	-.863
F40	.439	.559	.433	.350	.803	-.161

Raw Canonical Coefficients for Set-2

	1	2	3	4	5	6
F3	-.046	-.200	-.216	.204	.290	-.095
F8	.432	-.449	.130	.322	-.292	-.068
F12	-.039	-.148	.244	-.311	.235	.321
F18	.055	.049	-.418	-.066	-.124	.388
F27	.056	.232	-.148	-.423	-.221	-.435
F40	.305	.387	.300	.243	.557	-.111

Canonical Correlation

Canonical Loadings for Set-1

	1	2	3	4	5	6
F2	.553	.351	.322	.269	.214	-.528
F6	.576	.185	-.386	.470	.446	.209
F7	.779	-.541	.205	.155	-.171	-.012
F10	.628	-.023	-.003	-.139	.195	-.426
F14	.728	.075	.495	-.238	.387	.082
F20	.664	.254	-.377	-.462	.177	-.102
F26	.755	.480	.158	-.085	-.358	.047

Cross Loadings for Set-1

	1	2	3	4	5	6
F2	.409	.117	.086	.046	.027	-.023
F6	.426	.062	-.103	.081	.056	.009
F7	.576	-.180	.055	.027	-.022	-.001
F10	.465	-.008	-.001	-.024	.025	-.019
F14	.539	.025	.133	-.041	.049	.004
F20	.491	.085	-.101	-.080	.022	-.004
F26	.559	.160	.042	-.015	-.045	.002

Canonical Loadings for Set-2

	1	2	3	4	5	6
F3	.022	-.512	-.527	-.065	.590	-.328
F8	.856	-.483	.052	-.021	-.163	-.061
F12	.298	-.455	.163	-.688	.405	.202
F18	.575	.202	-.608	-.119	.024	.493
F27	.511	-.069	-.298	-.649	.120	-.458
F40	.742	.491	.045	.083	.445	.037

Cross Loadings for Set-2

	1	2	3	4	5	6
F3	.017	-.170	-.141	-.011	.075	-.014
F8	.633	-.161	.014	-.004	-.021	-.003
F12	.220	-.151	.044	-.118	.051	.009
F18	.426	.067	-.163	-.021	.003	.022
F27	.378	-.023	-.080	-.112	.015	-.020
F40	.549	.163	.012	.014	.056	.002

Note: Highlighted in green are those canonical loadings that are deemed significant based on the sample size of N=288 (0.35 or greater)

Canonical Correlation

Redundancy Analysis:

Proportion of Variance of Set-1 Explained by Its Own Can. Var.

	Prop Var
CV1-1	.455
CV1-2	.107
CV1-3	.101
CV1-4	.088
CV1-5	.089
CV1-6	.075

Proportion of Variance of Set-1 Explained by Opposite Can.Var.

	Prop Var
CV2-1	.249
CV2-2	.012
CV2-3	.007
CV2-4	.003
CV2-5	.001
CV2-6	.000

Proportion of Variance of Set-2 Explained by Its Own Can. Var.

	Prop Var
CV2-1	.328
CV2-2	.165
CV2-3	.128
CV2-4	.153
CV2-5	.125
CV2-6	.101

Proportion of Variance of Set-2 Explained by Opposite Can. Var.

	Prop Var
CV1-1	.179
CV1-2	.018
CV1-3	.009
CV1-4	.005
CV1-5	.002
CV1-6	.000

█ Indicates related to a statistically significant canonical correlation
----- END MATRIX -----

PART IV: TABLING RESULTS

SET 1

- F2 - Inner Harmony
- F6 - A Spiritual Life
- F7 - Sense of Belonging
- F10 - Meaning in Life
- F14 - Self Respect
- F20 - Self-Discipline
- F26 - Wisdom

SET 2

- F3 - Social Power
- F8 - Social Order
- F12 - Wealth
- F18 - Respect for Tradition
- F27 - Authority
- F40 - Honoring of Parents and Elders

Table 1. Significant Canonical Correlation #1:

	CV1-1 (_____)	CV2-1 (_____)
	45.5%	32.8%
	[24.9%]	[17.9%]
<u>Set 1</u>		
	<u>Loading</u>	<u>Loading</u>
F2	.553*	.022
F6	.576*	.856*
F7	.779*	.298
F10	.628*	.575*
F14	.728*	.511*
F20	.664*	.742*
F26	.755*	
	$R_c = 0.740$ $R_c^2 = 54.8\%$	<u>Set 2</u>
		F3
		F8
		F12
		F18
		F27
		F40
Wilk's Lambda	.356	
Chi-Square	209.639	
df	42	
Significance	$p < .001$	

Table 2. Significant Canonical Correlation #2:

	CV1-2 (_____)	CV2-2 (_____)	
	10.7%	16.5%	
	[1.2%]	[1.8%]	
<u>Set 1</u>	<u>Loading</u>		<u>Loading</u> <u>Set 2</u>
F2	.351*	$R_c = 0.333$ $R_c^2 = 11.1\%$	F3
F6	.185		F8
F7	-.541*		F12
F10	-.023		F18
F14	.075		F27
F20	.254		F40
F26	.480*		
Wilk's Lambda	.787		
Chi-Square	48.711		
df	30		
Significance	$p < .05$		

PART V: WRITEUP OF RESULTS

The results of the canonical correlation produced two significant canonical correlations, shown in Table 1 and Table 2. The first significant canonical correlation produced a Wilk's Lambda that was found to be highly significant through the use of a chi-square test that yielded a $p < .001$. The second canonical correlation also proved significant at the $p < 0.05$ level. All other canonical correlations produced by SPSS were found to be non-significant.

The first canonical correlation produced an R_c of 0.740, which indicates that the two variates have a shared variance of 54.8%, which can be found by squaring the R_c . In this correlation all seven (7) variables from SET 1 were found to have significant positive loadings on CV1-1, while four (4) of the six (6) variables in SET 2 were found to have significant positive loadings on CV2-1. Significance was determined by using Hair et al.'s factor loading guidelines. By this standard, the loadings were considered to be significant if they exceeded 0.35, based on the sample size of $n = 288$.

Due to CV1-1's strong showing by all seven variables in SET 1 it was given the title "self-actualization", reflecting upon the internal, positive values associated with the pool of SET 1 variables. The variables that loaded were:

- F2 – Inner Harmony
- F6 – A Spiritual Life
- F7 – Sense of Belonging
- F10 – Meaning in Life
- F14 – Self Respect
- F20 – Self-Discipline
- F26 – Wisdom

Meanwhile, CV2-1 had strong loadings by four of 6 variables in SET 2. Because of these loadings, CV2-1 was given the title of “Authoritarianism”. The variables that loaded are:

- F8 – Social Order
- F18 – Respect for Tradition
- F27 – Authority
- F40 – Honoring of Parents and Elders

The second canonical correlation produced a much smaller R_c of 0.333, with a corresponding shared variance between the variates of 11.1%. In this correlation, three (3) of the SET 1 variables showed significant loadings on CV1-2, with F7 (sense of belonging) showing a negative correlation, and the other two variables showing positive loadings. Due to the variables that were found to have loaded, the variate CV1-2 was titled “Hermiticism”. The variables that loaded are:

- F2 – Inner Harmony
- F7 – Sense of Belonging (NEGATIVE)
- F26 – Wisdom

For SET 2, four (4) variables showed significant loadings, of which three showed negative correlations. This led to the variate CV2-2 being titled “Tribalism”. The variables that loaded are:

- F3 – Social Power (NEGATIVE)
- F8 – Social Order (NEGATIVE)
- F12 – Wealth (NEGATIVE)
- F40 – Honoring of Parents and Elders

In summary, the more an individual places value on self-actualization variables like inner harmony and self-respect, the more likely they will also endorse more traditional, authoritarian values. Similarly, those who place a high value on a hermetic lifestyle are more likely to eschew money and power in favor of honoring kinship ties.