Chapter 5

Data Analysis and Results

5.1 Introduction

Chapter 5 begins with a discussion on data preparation that includes missing data, outliers, assessment of construct reliability and validity, and multicollinearity. Chapter 5 continues with a brief description of respondents' profile, followed by detailed results of analyses using a three-step clustering approach (Singh 1990). Clustering results are interpreted, with particular emphases on two cases of two-group comparisons. Chapter 5 ends with a presentation of hypotheses testing results.

5.2 Data Preparation

Data preparation involves such preliminary investigations as checking for missing data, removing outliers, assessing reliability and validity of measured constructs, and evaluating multicollinearity.

Prior to these activities, however, the very first task of data preparation concerns exclusion of cases that do not meet the scope of interest of the present study. As breast enlargement is the major concern of this study, respondents who stated that their desirable breast sizes are smaller than their current sizes were excluded (n = 69). Married respondents also were excluded (n = 59), for two reasons: To maintain homogeneity of the original target sample of the study and to avoid measurement error resulting from including married women who often have different consumption patterns from single women (Allen and Pickett 1987). On these bases, 125 cases were excluded from the data set and the remaining 590 cases were retained for further examination.

5.2.1 Missing Data

Of the 590 cases, three were deleted on the basis of missing data in past enlargement experience, a cluster validation variable of major interest to the present study. In addition, 44 cases were removed because of missing data in one of the two clustering variables, please oneself. A *t*-test was conducted to examine whether or not these missing data occurred at random. The data set was divided into two groups, with cases having missing data on please oneself in one group and those with valid responses on please oneself in the other group.

The two groups were compared with regard to the nine self-concept variables of central interest. As shown in Table 5.1, *t*-test results found no significant differences between the groups on the nine self-concept variables, except for physical vanity. Thus, it can be concluded that missing data occurred randomly rather than systematically.

Table 5.1 Assessing Randomness of Missing Data through Group Comparison of Cases with Missing versus Valid Data on Please Oneself

Self-Concept Variables	Cases with Missing Data		Cases with no Missing Data		*t*	*p**
	Mean	*n*	Mean	*n*		
Self-Esteem	4.52	44	4.50	545	0.17	0.866
Extraversion	3.92	44	4.01	545	-0.71	0.476
Openness to Experience	4.23	44	4.35	544	-0.92	0.357
Excitement	4.16	44	4.41	545	-1.88	0.060
Fun and Enjoyment	4.52	44	4.65	545	-0.88	0.377
Physical Vanity	3.77	44	4.24	545	-3.11	0.002
Achievement Vanity	4.17	44	4.25	544	-0.64	0.524
Being Well-Respected	4.02	44	4.17	545	-0.98	0.329
Self-Fulfillment	4.14	44	4.16	544	-0.15	0.884

*Two-tailed test

The removal of cases with missing data as discussed above reduced sample size to 543. These cases were subjected to an analysis of outliers, which will be discussed next.

5.2.2 Outliers

Many statistical procedures, including cluster analysis, are sensitive to outliers. Of particular concern to the present study is the possibility of outliers for the nine self-concept variables. Since these variables will be used later to describe each derived cluster, the presence of outliers will disrupt the description process. Therefore, the aim of outlier analysis is to detect cases that either are unique or extreme compared to the ordinary range of values on each of the nine self-concept variables or are unique or extreme in a combination of their values across two or more of the nine self-concept variables. Based on a review of values for each case on each of the nine self-concept variables, no univariate outliers could be identified at this point (in contrast to a univariate analysis of outliers conducted later within derived clusters, described in Section 5.7.1).

A multivariate assessment across the nine self-concept variables did detect outliers. Mahalanobis D^2 was used to measure the multidimensional position of each case relative to the column vector of means of the nine variables. Mahalanobis D^2 is a measure of distance in multidimensional space of each case to the column vector of means or centroid of the cases. It provides a common measure of multidimensional centrality relative to the centroid and has statistical properties that allow for significance testing. In the present study, Mahalanobis D^2 values were calculated based on the nine self-concept variables. A significance level of 0.01 (t = 2.33) was used as the threshold value for designation as an outlier. Since the D^2/df (where df denotes degrees of freedom and equals eight in this case) is approximately distributed as a t value (Hair Jr. *et al.* 1998), this translates into a threshold value for D^2 of 18.6. Cases with Mahalanobis D^2 exceeding a cutoff value of 18.6 were removed from the data set. Twenty-eight cases were removed on the basis of multivariate outliers, bringing the number of usable cases to 515.

5.2.3 Assessment of Construct Reliability and Validity

The assessment of construct reliability used Cronbach's alpha reliability coefficients. As discussed in Chapter 4, all constructs employed in the present study were reduced sets of established scales, with the number of scale items ranging from three to five. The number of scale items was kept small to comply with permitted survey time of approximately 25 minutes for each in-class survey (including time allocated for in-class instruction at the beginning of each survey).

Table 5.2 presents Cronbach's alpha coefficients for the nine self-concept variables, along with coefficient values reported from original studies that are sources of these constructs. With the exception of fun and enjoyment and being well-respected, Cronbach's coefficient values obtained in this study are lower than those reported in original studies. Among the many reasons for reductions in the magnitudes of reliability coefficients are the fewer number of items used in the present study and greater homogeneity in the sample. Although some of the present study's Cronbach's coefficient values are lower than the generally agreed upon lower limit of 0.7 (Peterson 1994), this should be acceptable given that coefficient alphas of the established scales achieve larger values because established scales contain a larger number of items. The acceptable levels of alphas indicate the ability of scale items to converge as measurements of their intended constructs, indicating convergent validity.

Table 5.2 Coefficient Alpha Values for the Nine Self-Concept Constructs (*n* = 515)

Construct	Present Study's Reliability Coefficients		Reported Reliability Coefficients		Source
	Coefficient Alpha	Number of Items	Coefficient Alpha	Number of Items	
Self-Esteem	0.56	4	0.77 – 0.88	10	Blascovich and Tomaka 1991
Extraversion	0.62	5	0.79	12	Costa and McCrae 1992a
Openness to Experience	0.59	4	0.80	12	Costa and McCrae 1992a
Excitement	0.69	3	0.72	4	Herche 1994
Fun and Enjoyment	0.79	3	0.79	4	Herche 1994
Physical Vanity	0.73	3	0.86	5	Netemeyer, Burton, and Lichtenstein 1995
Achievement Vanity	0.61	3	0.80	5	Netemeyer, Burton, and Lichtenstein 1995
Being Well-Respected	0.80	3	0.71	4	Herche 1994
Self-Fulfillment	0.58	3	0.77	5	Herche 1994

Table 5.3 presents the present study's item correlation matrix. In the matrix, higher correlations are expected between a scale item and other scale items belonging to the same construct than those between that scale item and items not belonging to that construct. For example, correlations between SE1 (self-esteem scale item 1) and SE2, SE3, and SE4 (self-esteem scale items 2, 3, and 4) are expected to be higher than correlations between SE1 and scale items other than SE2, SE3, and SE4. Visual inspection reveals that each scale item correlates relatively well with the other scale items belonging to the same construct. These correlations are relatively high compared with correlations between that same scale item and scale items of other constructs, indicating discriminant validity. The exception is EX3 (excitement scale item 3), which is highly correlated with FU1, FU2, and FU3 (fun and enjoyment scale items 1 to 3). This could lead to a multicollinearity problem when performing discriminant analysis when all nine self-concept variables will be included as predictor variables. The possible multicollinearity problem between excitement and fun and enjoyment will be addressed in the next section.

Lastly, Table 5.4 summarizes means, standard deviations, and the observed correlation matrix for the nine self-concept variables using the final data (31 items and n = 515). Each self-

concept construct is represented by a summated scale formed by combining its individual scale items into a single composite variable. Table 5.4 shows generally small positive correlations among the variables, a favorable result for multicollinearity and discriminant analysis to be described later.

Table 5.3 Item Correlation Matrix

	SE1	SE2	SE3	SE4	ET1	ET2	ET3	ET4	ET5	OP1	OP2	OP3	OP4	EX1	EX2	EX3	FU1	FU2	FU3	VP1	VP2	VP3	VC1	VC2	VC3	WR1	WR2	WR3	SF1	SF2	SF3
SE1																															
SE2	0.354**																														
SE3	0.147**	0.269**																													
SE4	0.086*	0.296**	0.316**																												
ET1	0.068	0.028	-0.067	-0.098*																											
ET2	0.194**	0.101*	0.065	0.027	0.257**																										
ET3	0.174**	0.103*	-0.014	0.020	0.362**	0.473**																									
ET4	-0.087*	0.037	0.000	0.044	0.165**	0.130**	0.167**																								
ET5	-0.031	0.124**	0.092*	0.105*	0.108*	0.074	0.138**	0.572**																							
OP1	0.114**	-0.008	-0.062	-0.028	0.005	0.188**	0.157**	-0.072	-0.049																						
OP2	0.125**	0.094*	0.077	-0.018	0.090*	0.137**	0.173**	0.065	0.025	0.261**																					
OP3	0.227**	0.123**	0.041	0.096*	0.070	0.107*	0.209**	-0.078	0.025	0.253**	0.379**																				
OP4	0.078	-0.059	-0.094*	0.017	-0.029	-0.057	0.061	-0.081	0.034	0.170**	0.190**	0.381**																			
EX1	0.156**	0.100*	-0.005	-0.028	0.186**	0.232**	0.275**	-0.004	-0.029	0.130**	0.295**	0.404**	0.181**																		
EX2	0.145**	0.079	-0.019	0.010	0.266**	0.216**	0.379**	0.055	0.016	0.141**	0.273**	0.349**	0.211**	0.577**																	
EX3	0.142**	0.085	0.032	-0.001	0.297**	0.325**	0.460**	0.142**	0.074	0.026	0.131**	0.133**	0.071	0.290**	0.415**																
FU1	0.137**	0.034	-0.050	-0.040	0.337**	0.330**	0.371**	0.132**	0.050	0.030	0.099*	0.159**	0.004	0.367**	0.411**	0.632**															
FU2	0.127**	0.041	-0.078	-0.052	0.276**	0.257**	0.295**	0.103*	0.005	0.003	0.112*	0.156**	0.040	0.362**	0.427**	0.546**	0.688**														
FU3	0.148**	0.134**	0.068	0.090*	0.202**	0.239**	0.363**	0.053	0.028	0.053	0.148**	0.171**	0.124**	0.316**	0.361**	0.555**	0.502**	0.481**													
VP1	0.114**	0.089*	-0.096*	0.022	0.283**	0.121**	0.203**	0.003	-0.031	0.072	0.089*	0.099*	0.070	0.150**	0.237**	0.262**	0.318**	0.304**	0.216**												
VP2	-0.056	-0.078	-0.208**	-0.137**	0.224**	0.048	0.163**	-0.004	-0.029	0.037	0.002	0.052	0.055	0.096*	0.110*	0.090*	0.170**	0.178**	0.027	0.526**											
VP3	0.083	0.063	-0.132**	0.024	0.214**	0.109*	0.195**	-0.031	0.011	0.139**	0.017	0.201**	0.051	0.214**	0.203**	0.215**	0.327**	0.299**	0.178**	0.493**	0.428**										
VA1	0.154**	0.051	-0.014	-0.088*	0.265**	0.100*	0.159**	-0.037	-0.033	0.045	0.077	0.198**	0.066	0.158**	0.177**	0.115**	0.224**	0.245**	0.192**	0.297**	0.204**	0.328**									
VA2	-0.012	-0.026	-0.200**	-0.034	0.033	0.010	-0.023	-0.109*	-0.076	-0.024	-0.024	0.076	0.089*	0.091*	0.045	-0.023	0.051	0.088*	0.023	0.238**	0.294**	0.292**	0.294**								
VA3	0.052	-0.025	-0.114**	-0.011	0.163**	0.014	0.024	-0.045	-0.017	-0.031	-0.019	0.112*	0.135**	0.116**	0.169**	-0.002	0.082	0.137**	-0.010	0.293**	0.255**	0.235**	0.235**	0.475**							
WR1	-0.056	-0.086	-0.225**	-0.141**	0.203**	0.070	0.073	-0.073	-0.117*	-0.003	-0.051	0.019	0.054	0.038	0.075	0.047	0.143**	0.156**	-0.026	0.364**	0.314**	0.226**	0.182**	0.287**	0.322**						
WR2	-0.075	-0.106*	-0.241**	-0.163**	0.246**	0.078	0.113*	0.081	-0.018	0.003	0.083	0.022	-0.009	0.084	0.148**	0.101*	0.211**	0.215**	0.044	0.257**	0.277**	0.156**	0.278**	0.227**	0.179**	0.520**					
WR3	-0.108*	-0.091*	-0.220**	-0.156**	0.232**	0.081	0.121**	0.057	-0.010	0.011	0.061	-0.010	-0.001	0.107*	0.130**	0.090*	0.190**	0.188**	0.055	0.330**	0.342**	0.236**	0.237**	0.209**	0.179**	0.493**	0.728**				
SF1	0.212**	0.193**	0.153**	0.174**	0.056	0.134**	0.109*	-0.010	0.020	0.062	0.121**	0.203**	0.103*	0.140**	0.262**	0.194**	0.172**	0.172**	0.169**	0.125**	-0.062	0.099*	0.054	-0.047	0.035	-0.009	0.046	0.021			
SF2	0.237**	0.161**	0.068	0.121**	0.157**	0.131**	0.117**	-0.032	-0.023	0.052	0.174**	0.267**	0.138**	0.316**	0.294**	0.236**	0.232**	0.315**	0.261**	0.226**	-0.016	0.213**	0.202**	0.100*	0.175**	0.086	0.054	0.042	0.406**		
SF3	0.175**	0.100*	-0.031	0.052	0.217**	0.062	0.191**	0.002	-0.016	0.043	0.122**	0.172**	0.145**	0.311**	0.305**	0.279**	0.316**	0.373**	0.219**	0.290**	0.155**	0.292**	0.170**	0.133**	0.207**	0.136**	0.109*	0.112*	0.154**	0.416**	

**Correlation is significant at 0.01 level (two-tailed)

*Correlation is significant at 0.05 level (two-tailed)

Note: SE denotes self-esteem, ET extraversion, OP openness to experience, EX excitement, FU fun and enjoyment, VP physical vanity, VA achievement vanity, WR being well-respected, and SF self-fulfillment.

Table 5.4 Correlation Matrix of Self-Concept Variables

	Number of Items	Mean	Standard Deviation	Correlation Matrix								
				1	2	3	4	5	6	7	8	9
1. Self-Esteem	4	4.51	0.78									
2. Extraversion	5	4.04	0.73	0.09*								
3. Openness to Experience	4	4.38	0.76	0.08	0.11*							
4. Excitement	3	4.45	0.79	0.09*	0.39**	0.36**						
5. Fun and Enjoyment	3	4.69	0.87	0.06	0.38**	0.16**	0.68**					
6. Physical Vanity	3	4.21	0.95	-0.09	0.19**	0.13**	0.27**	0.32**				
7. Achievement Vanity	3	4.27	0.84	-0.07	0.05	0.11*	0.15**	0.17**	0.44**			
8. Being Well-Respected	3	4.18	0.94	-0.26**	0.14**	0.03	0.14**	0.19**	0.41**	0.37**		
9. Self-Fulfillment	3	4.19	0.87	0.17**	0.15**	0.24**	0.44**	0.41**	0.28**	0.26**	0.13**	

**Significant at 0.01 level (two-tailed)
*Significant at 0.05 level (two-tailed)

5.2.4 Multicollinearity

Multicollinearity occurs when any independent variable is highly correlated with one or more other independent variables used in the same data analysis procedure. Multicollinearity is a source of concern when performing discriminant analysis, when all nine self-concept variables are entered simultaneously into the model as predictor variables. Highly collinear variables can distort results, making it difficult to identify self-concept variables that discriminate between *a priori* groups. That is, highly collinear variables can result in discriminant coefficients being artificially attenuated or inflated and in coefficients having signs opposite to that predicted by theory.

In the present study, the assessment of multicollinearity was undertaken in two steps: preliminary identification of collinearity using the correlation matrix for the predictor variables and assessment of pairwise and multiple variable collinearity using the tolerance value and its inverse, the variance inflation factor (VIF). As shown in Table 5.4, a preliminary examination of the correlation matrix for the nine self-concept variables reveals moderate to somewhat high correlations in the following pairs: excitement with fun and enjoyment (0.68); excitement with self-fulfillment (0.44); and fun and enjoyment with self-fulfillment (0.41).

Next, tolerance and VIF values were calculated to measure the impact of multiple variable collinearity. The tolerance value is defined as 1.00 minus the proportion of a variable's variance explained by the other independent variables. Thus, a high tolerance value indicates little collinearity while a low tolerance value approaching zero indicates that the variable is almost totally accounted for by the other independent variables. The variance inflation factor

is the reciprocal of the tolerance value; thus a small VIF value is preferred as indicative of low multiple variable collinearity among independent variables.

In the present study, tolerance and VIF values were calculated by performing a multiple regression using the nine self-concept variables as independent variables and please oneself as a dependent variable. The choice of the dependent variable is completely arbitrary and has no influence on any estimate of multicollinearity. As shown in Table 5.5, all tolerance values exceed a value of 0.1, a common cutoff threshold. Similarly, no VIF value exceeds 10.0. Both values indicate inconsequential collinearity. Therefore, it can be concluded that multicollinearity should not be a concern in performing multivariate discriminant analysis with the nine self-concept variables used as predictor variables. The correlation between excitement and fun and enjoyment is high but not to the point of creating a multicollinearity problem.

Table 5.5 Testing for Multicollinearity

Independent Variables	Tolerance	Variance Inflation Factor (VIF)
Self-Esteem	0.876	1.141
Extraversion	0.806	1.240
Openness to Experience	0.839	1.191
Excitement	0.444	2.251
Fun and Enjoyment	0.489	2.047
Physical Vanity	0.674	1.485
Achievement Vanity	0.739	1.352
Being Well-Respected	0.726	1.378
Self-Fulfillment	0.714	1.401

5.3 Respondents' Profile

Profiles of the study's 515 respondents appear in Table 5.6. Slightly more than half of the respondents are students of Thammasat University, about one-fourth are students of Chulalongkorn University, and approximately one-fifth are students of Srinakarinwirot University. Slightly more than half of all respondents are undergraduate students. Almost

half major in Business, about one-third major in Economics, and one-fifth major in Social Science. Almost half of the respondents are between 21 and 25 years of age and slightly more than one-fourth are in their late twenties. Only 12.4 percent of respondents are 20 or below, suggesting that most of the undergraduate students surveyed would be in their junior or senior years. Compared to the average Thai female population (*National Statistical Office* 2002), the sample is younger and better educated. Since married respondents have been excluded from the data set, all respondents are single with two cases reporting separation or divorce. In terms of income, slightly more than half of all respondents have monthly incomes of less than 15,000 baht.

Table 5.6 Profiles of Respondents (*n* = 515)

Characteristic	Number of Cases	Percent
University		
Chulalongkorn	125	24.3
Thammasat	285	55.3
Srinakarinwirot	105	20.4
Composition		
Undergraduate	286	55.5
Graduate	229	44.5
Major		
Economics	177	34.4
Business	233	45.2
Social Science	105	20.4
Age		
20 or below	64	12.4
21 to 25	254	49.3
26 to 30	141	27.4
31 to 35	39	7.6
36 to 40	13	2.5
More than 40	4	0.8
Marital Status		
Single	513	99.6
Seperated/Divorced	2	0.4
Monthly Income (Baht)		
Less than 15,000	275	53.6
15,000 to less than 30,000	165	32.2
30,000 to less than 45,000	48	9.4
45,000 to less than 60,000	16	3.1
60,000 to less than 75,000	7	1.4
75,000 to less than 90,000	0	0.0
90,000 or above	2	0.4

Table 5.7 presents mean values of the nine self-concept variables and the two clustering variables—ideal-actual discrepancy and please oneself—for respondents of the three universities surveyed in this study. Results of one-way ANOVA tests show no significant differences among respondents of these universities at the 0.05 level in terms of key variables.

Table 5.7 Analysis of Institutional Bias among the Three Universities Surveyed (*n* = 515)

Constructs		Chulalongkorn University	Srinakarinwirot University	Thammasat University
Self-Esteem	Mean	4.62	4.45	4.49
	Std. Deviation	0.79	0.71	0.81
	n	125	105	285
	p	0.186		
Extraversion	Mean	4.05	4.02	4.05
	Std. Deviation	0.69	0.75	0.74
	n	125	105	285
	p	0.955		
Openness to Experience	Mean	4.35	4.41	4.38
	Std. Deviation	0.77	0.82	0.74
	n	125	105	285
	p	0.830		
Excitement	Mean	4.42	4.50	4.45
	Std. Deviation	0.81	0.89	0.74
	n	125	105	285
	p	0.765		
Fun and Enjoyment	Mean	4.68	4.75	4.67
	Std. Deviation	0.92	0.91	0.84
	n	125	105	285
	p	0.711		
Physical Vanity	Mean	4.38	4.13	4.17
	Std. Deviation	0.93	1.01	0.94
	n	125	105	285
	p	0.086		
Achievement Vanity	Mean	4.26	4.30	4.27
	Std. Deviation	0.89	0.93	0.78
	n	125	105	285
	p	0.918		
Being Well-Respected	Mean	4.10	4.13	4.24
	Std. Deviation	0.97	0.95	0.91
	n	125	105	285
	p	0.306		
Self-Fulfillment	Mean	4.12	4.21	4.21
	Std. Deviation	0.88	0.90	0.85
	n	125	105	285
	p	0.619		
Ideal-Actual	Mean	1.66	1.50	1.63
Discrepancy	Std. Deviation	0.73	0.68	0.70
	n	125	105	285
	p	0.197		
Please Oneself	Mean	4.66	4.50	4.72
	Std. Deviation	1.48	1.56	1.35
	n	125	105	285
	p	0.432		

Note: Mean values of nine latent constructs are calculated based on a summation of scale items.

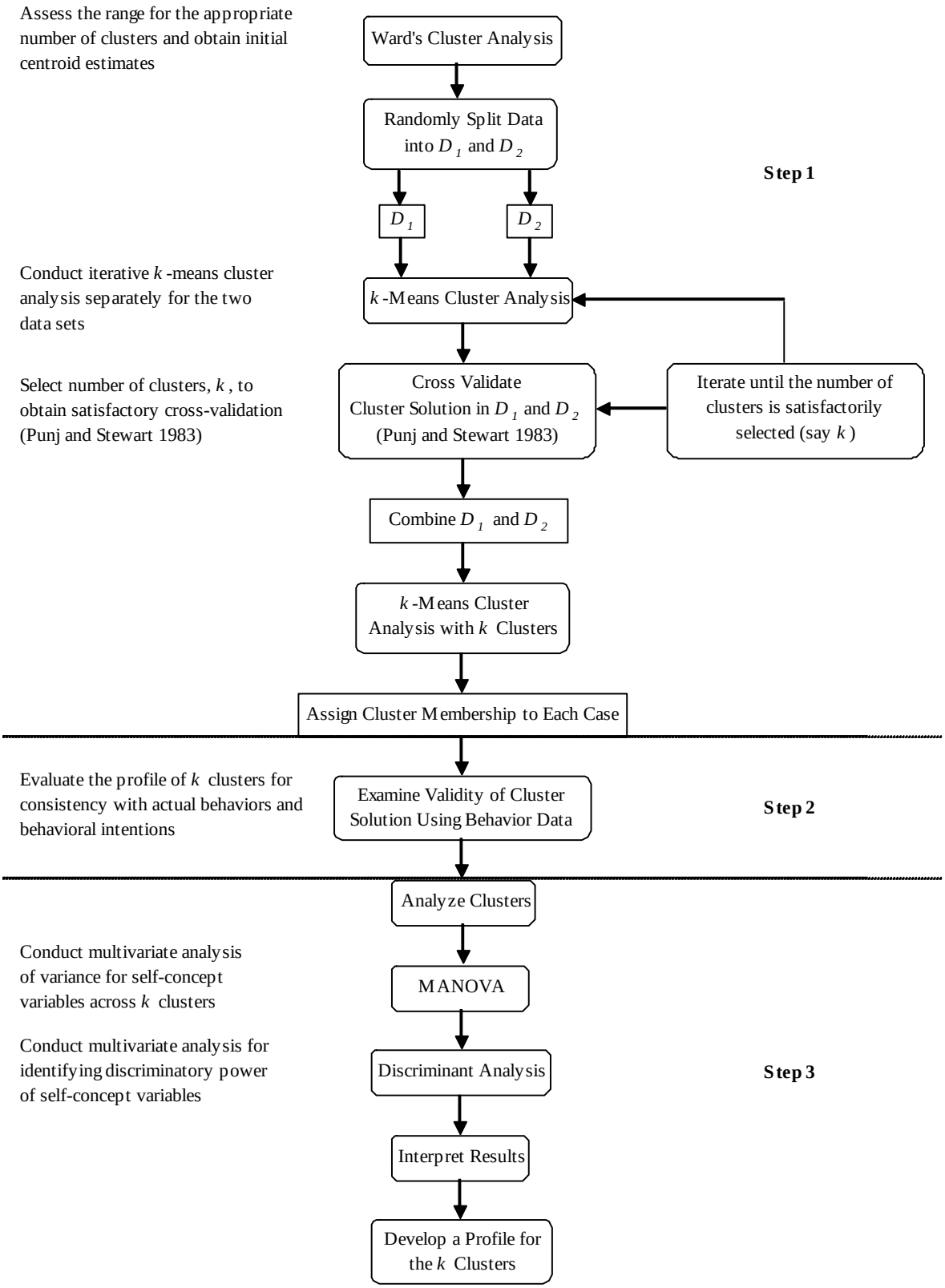
5.4 Overview of the Data Analysis Plan

A schematic diagram that overviews the method of analysis appears in Figure 5.1. As summary, cluster analysis is used to develop a profile of consumer segments using data driven procedures. Cluster analysis often produces no single best solution to any given clustering problem (Singh 1990) and provides no specified alpha level to guide selection of a particular solution from alternative solutions. Consequently, issues concerning validity and stability of cluster solutions become critically important (Punj and Stewart 1983; Singh 1990). Taking these validity and stability issues into account, the present study follows a three-step clustering approach suggested by Singh (1990).

In Step 1, the aim is to ascertain the optimal number of clusters (*k*) based on hierarchical and iterative cluster procedures and on internal validation of alternative cluster solutions. For the present cluster analysis, two variables—respondents' desirable incremental cup sizes compared to their current cup sizes (ideal-actual discrepancy) and the extent to which the decision to enlarge breasts is based on a desire to please oneself (please oneself)—were used as input to cluster analysis. These two variables are considered to be initial and central considerations in a woman's decision regarding breast enlargement. The two variables are measured with interval properties and allow the use of a distance metric as the basis to form clusters.

Following a procedure recommended by Singh (1990), the sample was randomly split into two data sets, D_1 and D_2, containing approximately 60 percent and 40 percent of the total number of cases respectively. D_1 was designated as the test sample and D_2 the internal validation sample. D_1 was utilized to generate possible alternative cluster solutions using a

Figure 5.1 The Three-Step Flow of Analysis Method Utilized in the Study



hierarchical clustering procedure. D_2 then was used to select the optimum non-hierarchical *k*-means clustering solution from several alternatives based on cluster stability and reproducibility. Stability and reproducibility were evaluated using the chance corrected coefficient of agreement, *Kappa*, computed for each selected value of *k* (Fleiss and Cohen 1973; Lorr 1983). The optimal *k* was chosen so as to maximize *Kappa*. Once the optimal *k* is determined, D_1 and D_2 were pooled and input into a final *k*-means cluster analysis with the number of clusters specified at the optimal value and the starting centroids specified at values obtained from Ward's method.

In Step 2, the aim is to establish external validation of the optimal cluster solution uncovered by *k*-means clustering. That is, derived *k*-means clusters necessarily will differ across clusters in terms of respondents' ideal-actual discrepancies and desires to please themselves. However, external validation of the optimal solution examines cluster differences in terms of actual behaviors and behavioral intentions. That is, to the extent that members of each derived cluster show between cluster differences in actual behaviors and behavioral intentions, the cluster solution would tend to be valid. Respondent's past behavior as to whether they had ever tried to enlarge breasts was used in this regard. Respondents' past enlargement behavior could include a wide variety of options available in the market, ranging from surgery, herbal creams, herbal pills, and contraceptive pills to alternatives that yield only temporary effects such as special bras or silicone pads. However, because surgery involves considerable financial and psychological risk, it was impossible to find respondents whose past enlargement behavior involved surgery. Consequently, instead of actual experience with surgery, the behavioral intention to undergo surgery was employed as the external validity check. For the *k* cluster solution, hypotheses are developed for variation in actual behaviors and behavioral intentions based on desirable incremental cup size and

desires of respondents to please themselves. These hypotheses are tested using contingency tables and chi-square statistics.

In Step 3, the aim is to distinguish among the derived *k* clusters in terms of the nine self-concept variables of interest to this study. First, MANOVA was used to assess the existence of overall differences among the *k* clusters on the entire set of self-concept variables. Discriminant analysis, an analysis method mathematically equivalent to MANOVA, then was employed to identify self-concept variables that have substantive discriminating power among the *k* clusters. The discriminating power of the total set as well as any subsets of self-concept variables are evaluated by Wilks' *Lambda* and variance explained, I^2. Wilks' *Lambda* provides a multivariate test of the discriminating power of the derived discriminant functions. I^2 is analogous to R^2 in multiple regression, measuring the amount of variance in the dependent variable explained by predictor variables acting together as a set. Significant discriminant functions then are interpreted on the basis of structure coefficients to identify self-concept variables that have substantive discriminating power. Details on the three-step clustering approach along with its results are discussed in the following sections.

5.5 Cluster Analysis

In Step 1 of the analysis, the aim is to ascertain the optimal number of clusters (*k*) based on hierarchical and non-hierarchical cluster analysis and internal validation of alternative cluster solutions. To describe the first step, this section begins with a brief introduction to the study's clustering variables followed by a more complete discussion of hierarchical and non-hierarchical cluster analyses. This section ends by presenting results obtained from cluster analysis.

5.5.1 Clustering Variables

The study's two clustering variables consist of respondents' desirable incremental cup sizes compared with their current cup sizes (ideal-actual discrepancy) and the extent to which the decision to enlarge breasts is based on a desire to please oneself (please oneself). The correlation between the two clustering variables is small at 0.15 ($p < 0.01$), indicating that each variable provides near-unique information about respondents. A formal assessment of multicollinearity (conducted by performing a multiple regression using ideal-actual discrepancy and please oneself as independent variables and physical vanity as the dependent variable) revealed that both clustering variables have a tolerance and VIF values of 0.977 and 1.023, respectively. Both values indicate that multicollinearity is not a concern to performing cluster analysis.

5.5.2 Hierarchical Cluster Analysis

Objectives of hierarchical cluster analysis are threefold: to detect outliers that are unacceptably remote in terms of mean squared distance from other cases in the same cluster with respect to the clustering variables; to generate possible alternative cluster solutions to the classification problem; and to obtain a cluster seed for each cluster as *a priori* information for non-hierarchical *k*-means cluster analysis performed later.

At the beginning stage of hierarchical cluster analysis, two major decisions must be made in terms of what clustering method should be used to form clusters and how many clusters should be formed. As one of the study's clustering variables, ideal-actual discrepancy, involves four categorical levels—three-cup larger, two-cup larger, one-cup larger, and no change preferred, a minimum of four clusters is considered acceptable. On the other hand, if the other clustering variable, please oneself, is considered as well and is classified into three

categories—high, medium, and low—as an extreme scenario, a possible upper limit would reach 12 clusters. Based on these considerations, a range of 4 to 12 clusters was deemed appropriate.

Ward's method was chosen as the hierarchical method of clustering for this study. At each stage in the clustering procedure, the within-cluster sum of squares on the two clustering variables is minimized over all partitions (the complete set of disjoint or separate clusters) obtainable by combining two clusters from the previous stage. Although the method is biased toward production of clusters with approximately the same number of observations, this is considered appropriate given statistical power attending the resulting cluster solution when used in multiple-group analysis.

Next, an analysis of agglomeration coefficients produced by Ward's method is conducted in order to determine the optimal cluster solution range. Conceptually, a large increase in the agglomeration coefficient between any *k* and a *k+1* cluster solution indicates the merger of two very different clusters. Therefore, changes in the agglomeration coefficient serve to indicate the optimal number of clusters. According to an analysis of agglomeration coefficients for the Ward's method cluster analysis presented in Table 5.8, the agglomeration coefficient showed large increases when moving from two clusters to one (61.8 percent), from three clusters to two (23.4 percent), from four to three (10.3 percent), from five to four (9.7 percent), from six to five (12.3 percent), from seven to six (10.9 percent), and from eight to seven (11.3 percent). On this basis, the optimal cluster solution range was adjusted to between two and eight clusters. However, given the fact that a percentage change in agglomeration coefficient when moving from two clusters to one is usually large, the optimal cluster solution was finally adjusted to a range of three to eight clusters.

Table 5.8 Analysis of Agglomeration Coefficient for Hierarchical Cluster Analysis (*n* = 515)

Number of Clusters	Agglomeration Coefficient	Change in Agglomeration Coefficient to Next Level	Percentage Change in Coefficient to Next Level
12	111.55	6.47	5.8%
11	118.02	3.78	3.2%
10	121.79	7.67	6.3%
9	129.46	3.88	3.0%
8	133.35	15.07	11.3%
7	148.42	16.18	10.9%
6	164.59	20.25	12.3%
5	184.84	17.93	9.7%
4	202.77	20.89	10.3%
3	223.65	52.34	23.4%
2	275.99	170.56	61.8%
1	446.55		

Thus, alternative cluster solutions ranging from three to eight clusters were identified along with column vectors of means associated with each cluster in each solution. These column vectors will be used as starting points to initiate *k*-means clustering procedures, in which clusters are built around these pre-specified points (see Appendix 3). Results of the final hierarchical clustering in terms of cluster sizes are shown in Table 5.9 for the three to eight cluster solutions. Frequency distributions of the two clustering variables for each cluster solution indicated that no outliers exist as targets for deletion.

Table 5.9 Cluster Sizes for Final Hierarchical Clustering Results, Three to Eight Cluster Solutions (*n* = 515)

Cluster	Number of Cases for Each Cluster Solution					
	3	4	5	6	7	8
1	106	106	106	64	64	64
2	221	124	124	124	67	67
3	188	188	59	59	59	59
4		97	97	97	57	57
5			129	42	97	97
6				129	42	42
7					129	64
8						65

As the range of three to eight alternative cluster solutions have been identified from hierarchical cluster analysis, the next step proceeds to derive and internally validate a chosen cluster solution using *k*-means clustering.

5.5.3 *k*-means Cluster Analysis

The last procedure in Step 1 uses *k*-means clustering, a non-hierarchical procedure. The *k*-means cluster procedure serves two purposes: to adjust or "fine-tune" results obtained from hierarchical procedures and to internally validate the chosen cluster solution.

As mentioned earlier, the 515 cases were randomly split into two data sets, D_1 and D_2, containing 280 and 235 cases respectively. D_1 was designated as the test sample and D_2 the internal validation sample. The test sample used six initial centroids from Ward's hierarchical clustering from the three to eight cluster solutions to drive six "constrained" scenarios for *k*-means clustering of the test sample. Column vectors of means associated with each cluster for the range of three to eight cluster solutions produced by non-hierarchical *k*-means clustering on the D_1 test sample are provided in Appendix 4.

Results then were used to select the optimal solution among possible alternatives based on cluster stability and reproducibility using two scenarios for the D_2 internal validation sample: constrained and unconstrained. The constrained scenario classifies all cases in D_2 based on analysis results from the test sample. The proximity approach based on Euclidean distances among cases was applied to assign each case to its respective cluster. By contrast, the unconstrained solution posed no restrictions. Table 5.10 provides numbers of cases belonging to each cluster solution under both constrained and unconstrained scenarios for the D_2 sample.

Table 5.10 Summary of Number of Cluster Members for D_2 Constrained and Unconstrained Scenarios (n = 235)

Cluster	Number of Cases for Each Cluster Solution											
	3		4		5		6		7		8	
	C	U	C	U	C	U	C	U	C	U	C	U
1	55	40	39	40	39	40	32	35	26	28	22	25
2	97	97	70	76	70	70	54	54	23	23	23	23
3	83	98	83	76	27	26	27	24	26	24	26	24
4			43	43	43	43	43	43	38	38	30	27
5					56	56	23	23	43	43	23	18
6							56	56	23	23	23	18
7									56	56	32	44
8											56	56

Note: **C** denotes constrained and **U** unconstrained scenario.

Next, the chance corrected coefficient of agreement, *Kappa*, is computed for the two scenarios—constrained and unconstrained—of D_2 cases for each cluster solution. Table 5.11 presents results of *Kappa* along with other similar symmetric measures—*Gamma*, *Tau-B*, and *Tau-C*. *Kappa* tests whether agreement between constrained and unconstrained scenarios exceeds chance levels. The maximum value of *Kappa* identifies which cluster solution is the most stable under both constrained and unconstrained scenarios. The optimal cluster solution is chosen so as to maximize *Kappa*. Based on this criterion, the optimal cluster solution selected is the five-cluster solution.

Table 5.11 Cluster Solution and Chance Corrected Coefficients of Agreement (n = 235)

Cluster Solution	Symmetric Measures[a]			
	Kappa	*Gamma*	*Tau-B*	*Tau-C*
3	0.87	0.80	0.74	0.71
4	0.94	0.98	0.94	0.91
5	0.97	0.98	0.97	0.94
6	0.95	0.95	0.94	0.92
7	0.96	0.95	0.94	0.92
8	0.95	0.94	0.92	0.91

[a]All chi-square statistics are significant at 0.001 level

Once the optimal five-cluster solution was determined, the data (D_1 and D_2) were pooled and input into a final *k*-means cluster analysis to complete step one of the procedure. These final clustering results are presented next.

Table 5.12 contains mean values of clustering variables—ideal-actual discrepancy and please oneself—for the final *k*-means cluster solution, along with each cluster's respective number of cases. Note that cluster mean values, represented here as *IAD* for ideal-actual discrepancy and as *PO* for please oneself, are based on raw data scores. Made up of nearly half of the respondents, Cluster 5 (*IAD* = 1.29, *PO* = 5.74), the largest cluster, can be described as respondents who view breast enlargement to please oneself as strongly important but who are almost satisfied with their current breast size. Cluster 3 (*IAD* = 3.12, *PO* = 6.00), the smallest cluster, can be characterized as respondents who have the largest ideal-actual discrepancy toward their own current breast sizes in comparison to all other clusters and view breast enlargement to please oneself as strongly important. Cluster 1 (*IAD* = 1.22, *PO* = 3.48) contains respondents who are satisfied with their current breast sizes and are neutral in terms of perception toward breast enlargement as a way to please oneself. Cluster 2 (*IAD* = 1.31, *PO* = 1.62) contains respondents who are satisfied with their current breast sizes and view breast enlargement to please oneself as almost not at all important. Cluster 4 (*IAD* = 2.21, *PO* = 4.64), the second largest cluster, contains respondents who have moderate ideal-actual discrepancy toward their own current breast sizes and moderately favor breast enlargement as a way to please oneself.

Table 5.12 Mean Values of Clustering Variables (*n* = 515)

Cluster	Ideal-Actual Discrepancy[a] (*IAD*)	Please Oneself[b] (*PO*)	Number of Cases	Percentage of Respondents
1	1.22	3.48	86	16.7
2	1.31	1.62	55	10.7
3	3.12	6.00	28	5.4
4	2.21	4.64	128	24.9
5	1.29	5.74	218	42.3
Mean	1.60	4.66		

[a]Based on raw data scores. 1 signifies just right; 2 one cup size larger; 3 two cup size larger; and 4 three cup size larger.
[b]Based on raw data scores. 1 signifies not at all important; 2 quite unimportant; 3 somewhat unimportant; 4 somewhat important; 5 quite important; and, 6 strongly important.

5.6 External Validity Test

Step 2 in Singh's (1990) procedure subjects the optimal cluster solution uncovered by *k*-means clustering to external validation. Given the final *k*-means cluster solution, actual behavior and behavioral intention in a practical setting can be used to assess cluster validity. If expectations are confirmed, it follows that external validity of the procedure and solution are supported, suggesting that the derived cluster solution is practically valid.

As a preparatory step for the test of external validity, a multiplicative composite variable, *IADPO*, was created as:

$$IADPO = IAD \times PO \text{ ,}$$

IADPO denotes a composite variable that takes into account the effect of both ideal-actual discrepancy (desirable incremental cup sizes) and the extent to which the decision to enlarge breasts is based on the desire to please oneself. *IAD* represents ideal-actual discrepancy score (desirable incremental cup sizes score) and *PO* signifies the score for the extent to which the decision to enlarge breasts is based on a desire to please oneself. Based on this definition, the largest *IADPO* is observed in Cluster 3 (18.7), followed by Cluster 4 (10.3), Cluster 5 (7.4), Cluster 1 (4.3), and Cluster 2 (2.1), respectively.

The test of external validity was conducted on two variables. The first is past enlargement experience, an actual behavioral variable measured on a dichotomous basis (Yes/No). It is expected that variability of past enlargement experience should be observed across the five clusters. That is, for the cluster solution to be valid, past enlargement experience should be most frequently observed in the cluster whose members perceive their ideal breast size to be substantially larger than their current size while, at the same time, viewing pleasing themselves as a strongly important basis for embarking on breast enlargement. The

experience should be least frequently found among those who feel satisfied with their current breast size and do not perceive pleasing themselves as an important basis for enlargement. In other words, the cluster with the highest value of *IADPO* should witness the largest proportion of respondents with past enlargement experience, followed by the cluster with the next highest value of *IADPO*, and so on.

The other variable used in the test for external validity is intention to undergo surgery. That is, instead of actual experience, the behavioral intention to undergo surgery was employed as a proxy for the validating behavior. It is expected that proportions of respondents who would choose surgery as an augmentation method will vary across clusters, with the highest proportion found in the cluster having the highest *IADPO*.

Table 5.13 presents results for the analysis of external validity. Variation across the five clusters is significant for past enlargement experience ($\chi^2 = 23.6$, $p < 0.001$) and the expected pattern is strongly supported with one exception. Past enlargement experience in terms of percentage of respondents is largest in Cluster 3 (the cluster with the highest value of *IADPO*) at 46.4 percent. As expected, Cluster 4 has the next highest proportion of respondents with past enlargement experience, followed by Cluster 5. However, contrary to expectations, Cluster 1 has a lower proportion of those with past enlargement experience compared to Cluster 2. However, this difference is very small, with 15.1 percent for Cluster 1 versus 16.4 percent for Cluster 2.

Table 5.13 Validity Check for the Cluster Solution (*n* = 515)

Validity Check Variables	Cluster					Chi-Square Value (Probability)[a]
	1	2	3	4	5	
Past Enlargement Experience	13	9	13	45	44	23.6 (0.000)
Number of Cluster Members	86	55	28	128	218	
Percent of Respondents	15.1	16.4	46.4	35.2	20.2	
Intention to Undergo Surgery[b]	12	18	15	35	65	14.5 (0.006)
Number of Cluster Members[b]	258	171	84	387	654	
Percent of Respondents	4.7	10.5	17.9	9.0	9.9	

[a]Tests the null hypothesis that the validity check variables (e.g., past enlargement experience) are independent of cluster membership.
[b]Counts are based on three scenarios.

As for intention to undergo surgery, variation across the five clusters again is significant (χ^2 = 14.51, $p < 0.01$). As expected, intention to undergo surgery is most frequently observed in terms of percentage of respondents in Cluster 3, the cluster with the highest value of *IADPO*. Intentions to undergo surgery for Clusters 2, 4, and 5 are approximately half that of Cluster 3; intentions for Cluster 1 are approximately one-fourth that of Cluster 3.

Based on analyses described in this section, the derived five-cluster *k*-means solution can be considered externally validated. Attention moves in the next section to how this solution can be described in terms of the clustering variables and the self-concept variables.

5.7 Analyses and Interpretations of Derived Cluster Solution

Step 3 in Singh's (1990) procedure identifies characteristics that differ across the derived clusters with an ultimate aim of developing a profile that describes these clusters. The major concern for this step, therefore, is: How can the derived five-cluster solution be described in terms of self-concept variables? To accomplish this, the section begins by testing whether the

five clusters are different with respect to a linear combination of the nine self-concept variables. It then identifies the combination of self-concept variables that has optimum discriminating power among the derived clusters. The section ends by describing the five clusters using results of a series of analyses conducted in the section.

5.7.1 Analyses of Derived Cluster Solution

Prior to performing all analyses, a scatter plot identified three cases in three clusters as outliers. These cases had extreme values on a self-concept variable relative to mean values on that variable for all cases in the respective clusters. The first case belonged to Cluster 3 and was associated with openness to experience. The second case belonged to Cluster 4 and was associated with physical vanity. The third case belonged to Cluster 3 and was associated with being well-respected. These three cases were removed because, on each respective variable identified above, cases possessed a value smaller than three standard deviations from the mean compared to other members of the same cluster. These three cases were removed, bringing the number of usable cases in these analyses to 512.

Analysis of the derived five-cluster solution began with multivariate analysis of variance, MANOVA, to examine distinctions among clusters in terms of the self-concept variables of interest. Specifically, MANOVA was used to assess whether an overall difference exists among clusters on the nine self-concept variables. MANOVA detects cluster differences with respect to a linear combination of self-concept variables. MANOVA takes into account correlations among the self-concept variables and utilizes the total information available for assessing overall group differences that is missing when examining each self-concept dependent variable separately. The following test statistics were obtained from the

MANOVA test: Wilks' *Lambda* = 0.901; $F = 1.470$; and, $p < 0.036$. The five clusters as a whole exhibit significant differences on the set of self-concept variables.

The question now to be asked is: Which self-concept variables have discriminating power in differentiating among the five clusters? Several analyses can be conducted to answer this question. Among them are a series of univariate ANOVAs conducted on each self-concept variable. However, separate univariate ANOVAs ignore the total information available for evaluating overall cluster differences provided by MANOVA. Discriminant analysis is deemed more appropriate in this regard in that it provides information useful in assessing all nine self-concept variables simultaneously. Discriminant analysis produces a structure loading matrix that contains correlations between the linear combination of self-concept variables that maximizes cluster differences and the self-concept variables themselves. Self-concept variables that correlate strongly with the linear combination are more important in discriminating between clusters than self-concept variables that correlate weakly.

Discriminant analysis derives a variate or linear combination of two or more independent variables that discriminates best between *a priori* defined groups. Discrimination is achieved by setting the variate's weights for each variable that maximize between-group variance relative to within-group variance. As noted earlier, discriminant analysis is mathematically equivalent to MANOVA, yielding exactly the same test statistics as shown above. The variate or linear combination for a discriminant analysis, also known as the discriminant function, is an equation of the following form:

$$D_{jk} = a + W_{j1}X_{1k} + W_{j2}X_{2k} + ... + W_{jn}X_{nk}$$, where

D_{jk} = discriminant score of discriminant function j for individual k

a = intercept

W_{ji} = discriminant coefficient for discriminant function j and independent variable i

X_{ik} = independent variable i for observation k

In a standardized form, discriminant function takes a similar form:

$$Z_{jk} = d_{j1}X_{1k} + d_{j2}X_{2k} + ... + d_{jn}X_{nk}$$, where

Z_{jk} = standardized discriminant score of discriminant function j for individual k

d_{ji} = standardized discriminant function coefficient for discriminant function j and independent variable i

X_{ik} = standardized independent variable i for observation k.

The objective of discriminant analysis here is to determine which combination of self-concept variables has the best discriminating capabilities in discriminating among the five clusters. To achieve this goal, a discriminant analysis was performed simultaneously using the nine self-concept variables as predictors of group membership for the five derived clusters.

As a preparatory step, it is important to ensure that the data meet assumptions required by discriminant analysis: freedom from multicollinearity, linear relationships among all pairs of predictors, and variance/covariance equality among the groups. Earlier results in Subsection 5.2.4 demonstrated that the data are free from multicollinearity. As shown in Table 5.14, the maximum absolute values for skewness and kurtosis for the data are 0.52 and 0.45, lower

than the general rule of thumb of 0.75 and 1.50. The assumption of linearity is less serious and violations lead only to reduced power (Tabachnick and Fidell 2001, p. 463), not a concern in this study because of the large sample size. As a precaution, however, tests of linearity for the seven largest and seven smallest correlations in Table 5.4 showed only one small, significant departure from a linear relationship. To examine if the variance/covariance equality assumption is violated, the Box's *M* test statistic is evaluated to test the null hypothesis of equality of variance/covariance matrices across the five groups. The following statistics are obtained: Box's $M = 188.08$; approx. $F = 0.975$; $df = 180,46493$; $p = 0.58$. Since the null hypothesis cannot be rejected, the variance/covariance equality assumption is met.

Table 5.14 Analysis of Skewness and Kurtosis (*n* = 512)

	Mean	Standard Deviation	Skewness	Standard Error	Kurtosis	Standard Error
Self-Esteem	4.51	0.78	-0.36	.108	-0.11	.215
Extraversion	4.04	0.73	-0.03	.108	-0.24	.215
Openness to Experience	4.38	0.75	-0.37	.108	-0.03	.215
Excitement	4.45	0.79	-0.25	.108	-0.16	.215
Fun and Enjoyment	4.69	0.87	-0.52	.108	-0.02	.215
Physical Vanity	4.22	0.95	-0.18	.108	-0.23	.215
Achievement Vanity	4.28	0.84	-0.10	.108	-0.45	.215
Being Well-Respected	4.19	0.93	-0.33	.108	-0.14	.215
Self-Fulfillment	4.19	0.87	-0.17	.108	0.03	.215

Table 5.15 presents multivariate results for the five-cluster discriminant analysis. Of the four functions, only Function 1 is statistically significant as measured by the chi-square statistic. The first discriminant function always accounts for the largest amount of variation in the discriminant groups. The second discriminant function is orthogonal to the first and explains the largest percentage of variance remaining (after variance for the first function is removed). As shown in Table 5.15, the first function accounts for 58.8 percent of the variance explained by the four functions. The total amount of variance explained by the first function is $(0.243)^2$, or 5.9 percent. The next function explains $(0.161)^2$, or 2.6 percent, of the remaining variance (94.1 percent). Therefore, the total variance explained by Functions 1 and 2 is 5.9 percent +

(2.6 percent × 94.1 percent), or 8.3 percent of the total variation in the dependent variable. Following similar calculations, the total variance explained by Functions 1, 2, 3, and 4 is 10.0 percent. However, as Function 1 is the only statistically significant function, the present study largely limits its interpretation to that function.

Table 5.15 Multivariate Results for Five-Cluster Discriminant Analysis

Canonical Discriminant Functions

Function	Eigenvalue	Percent of Variance		Canonical Correlation	Percent of Total Variance Explained	
		Function	Cumulative		Function	Cumulative
1	0.063	58.8	58.8	0.243	5.9	5.9
2	0.027	25.0	83.8	0.161	2.4	8.3
3	0.014	13.4	97.2	0.119	1.4	9.7
4	0.004	2.8	100.0	0.055	0.3	10.0

Wilks' *Lambda*

Test of Function(s)	Wilks' *Lambda*	Chi-square	*df*	*p*
1 through 4	0.901	52.646	36	0.036
2 through 4	0.957	21.971	24	0.581
3 through 4	0.983	8.672	14	0.851
4	0.997	1.525	6	0.958

5.7.2 Interpretations of the Derived Cluster Solution

After estimating the discriminant functions, the next phase is interpretation. This stage involves examining discriminant function results to determine the relative importance of each self-concept variable in discriminating among clusters. Because structure loadings are considered more informative than standardized coefficients (Hair Jr. *et al.* 1998, p. 293; Tabachnick and Fidell 2001, p. 484), structure loadings are used for interpretation. Structure loadings are pooled within-groups correlations between the nine self-concept variables and standardized discriminant functions.

Table 5.16 contains structure loadings for the four discriminant functions. As there is no test of significance for structure loadings, simple correlations between each variable and the discriminant *Z* scores associated with the respective discriminant functions were computed

with results shown also in Table 5.16. With the exception of openness to experience in Function 3, all structure loadings represent significant correlations with their respective discriminant functions ($p < 0.01$). For interpretation purposes, independent variables are ranked in terms of their discriminating value in Table 5.16 based on their structure loadings. Signs associated with loadings do not affect the rankings; they simply indicate a positive or negative relationship between a self-concept variable with the indicated function.

Table 5.16 Summary of Interpretive Measures for Five-Cluster Discriminant Analysis (n = 512)

Independent Variables	Structure Loadings and Correlation Coefficients of Function 1			Structure Loadings and Correlation Coefficients of Function 2		
	Structure Loading	Rank	Simple[a] Correlation	Structure Loading	Rank	Simple[a] Correlation
Self-Esteem	0.236	9	0.241	0.750*	1	0.753
Extraversion	0.496*	3	0.507	0.194	8	0.195
Openness to Experience	0.301	6	0.308	-0.377	2	-0.380
Excitement	0.364*	5	0.373	0.286	4	0.288
Fun and Enjoyment	0.505*	2	0.515	0.248	5	0.249
Physical Vanity	0.850*	1	0.856	-0.255	6	-0.252
Achievement Vanity	0.294	7	0.302	-0.162	9	-0.164
Being Well-Respected	0.252	8	0.258	-0.313	3	-0.315
Self-Fulfillment	0.369	4	0.379	0.230	6	0.232

Independent Variables	Structure Loadings and Correlation Coefficients of Function 3			Structure Loadings and Correlation Coefficients of Function 4		
	Structure Loading	Rank	Simple[a] Correlation	Structure Loading	Rank	Simple[a] Correlation
Self-Esteem	0.297	5	0.297	0.234	5	0.232
Extraversion	-0.381	2	-0.380	0.215	7	0.214
Openness to Experience	0.049	9	0.049	0.536*	1	0.534
Excitement	-0.225	7	-0.225	0.268	4	0.267
Fun and Enjoyment	-0.476	1	-0.475	-0.227	6	-0.225
Physical Vanity	0.161	8	0.158	-0.129	9	-0.126
Achievement Vanity	0.264	6	0.265	0.315*	2	0.314
Being Well-Respected	0.341*	4	0.342	-0.308	3	-0.307
Self-Fulfillment	0.372*	3	0.372	-0.133	8	-0.133

*Largest absolute correlation between each variable and any discriminant function.
[a]Significant at $p < 0.01$ level except $p = 0.269$ for Openness to Experience in Function 3.

Which self-concept variables are substantive discriminators? There are no established criteria to answer this question and the present study applies four general rules of thumb discussed in the literature. First, for any variable to be qualified as a substantive discriminator on any

discriminant function, Wilks' *Lambda*, the measure of the discriminatory power of the discriminant function, must be statistically significant. More specifically, Wilks' *Lambda* at least should be significant on the first discriminant function, as this function extracts the most variance. Second, to be a substantive discriminator, the variable of interest must exhibit a structure loading of ±0.30 or higher (Hair Jr. *et al.* 1998; Tabachnick and Fidell 2001). Third, if more than one discriminant function is significant, the largest absolute correlation of the variable of interest with discriminant scores is used to identify substantive discriminators. When compared across discriminant functions, the largest absolute correlation indicates where a variable has the most discriminating power when included in the discriminant function under consideration. Fourth, to be a substantive discriminator, the variable must correlate significantly with discriminant scores for the discriminant function under consideration. The measure of correlation of the variable with the discriminant scores serves as an indicator for determining the statistical significance of the variable as a predictor variable in the discriminant model.

Based on the first rule discussed above, the present study should focus its interpretation largely on Function 1, as it is the only significant discriminant function (see Table 5.15). Applying the ±0.30-or-higher rule, extraversion, openness to experience, excitement, fun and enjoyment, physical vanity, and self-fulfillment qualify as substantive discriminators (see Table 5.16). In a less strict sense, achievement vanity also should qualify as its structure loading is very close to the 0.30 cutoff level.

Also, it should be emphasized that while extraversion, openness to experience, excitement, fun and enjoyment, physical vanity, achievement vanity, and self-fulfillment are highlighted, correlations of all self-concept variables with all discriminant scores for Function 1 are

significant at $p < 0.01$. This suggests that although their degrees of contribution differ, all nine self-concept variables as a combination play an important role in discriminating among the five clusters.

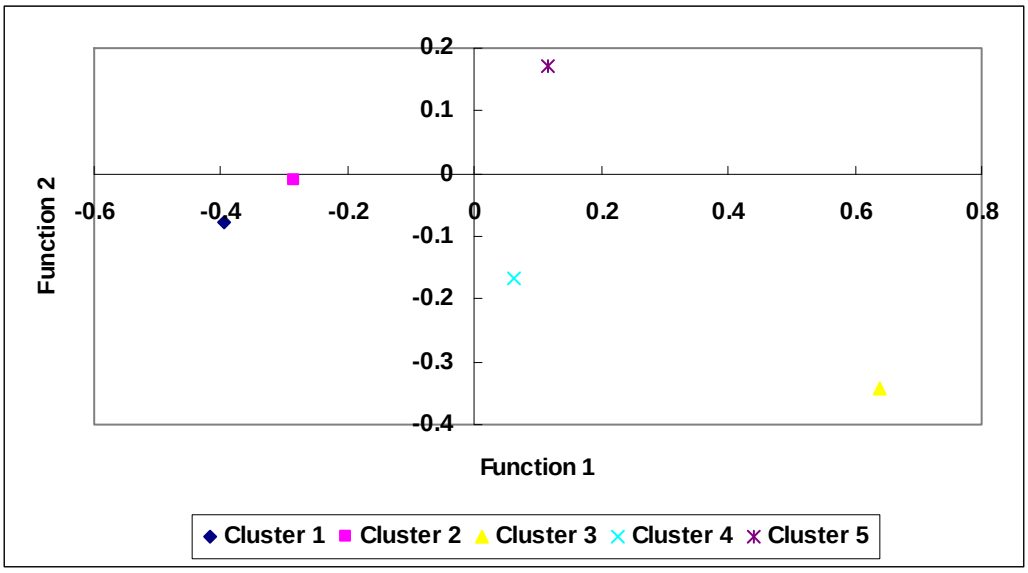
Additional interpretations of the above results can be conducted by a graphical illustration of cluster centroids in discriminant space and by judgment through the use of cluster centroids. Both approaches are discussed next.

To depict differences in terms of the predictor variables, cluster centroids can be plotted in discriminant function space in which the first two discriminant functions form horizontal and vertical axis. Interpretation of the plot in Figure 5.2 indicates that Discriminant Function 1 is the primary source of difference between: Clusters 1 and 2 versus Clusters 4 and 5; Clusters 1 and 2 versus Cluster 3; and, Clusters 4 and 5 versus Cluster 3. Moreover, Function 1 corresponds most closely to physical vanity, fun and enjoyment, extraversion, self-fulfillment, excitement, openness to experience, and achievement vanity in order of their absolute correlations of structure loadings with Function 1. Thus, the distinguishing characteristics of Cluster 3 are remarkably positive tendencies toward these seven self-concept variables.

Discriminant Function 2 provides the distinction between Cluster 3 versus Cluster 5, between Cluster 4 versus Cluster 5, and between Cluster 3 versus Clusters 1 and 2. The high absolute correlation of structure loading for self-esteem with Discriminant Function 2 signifies that self-esteem is most descriptive of the second discriminant function. As stated in the first rule of thumb, for any variable to be qualified as a substantive discriminator on a discriminant function, the discriminant function must be statistically significant as measured in terms of Wilks' *Lambda*. However, due to the fact that Discriminant Function 2 is not statistically

significant and that Function 2 represents substantially less variance than Function 1, caution should be observed in determining the impact of variables based on their loadings on this second function.

Figure 5.2 Cluster Centroids in Discriminant Function Space



The other approach to interpretation is to use structure loadings in Table 5.16 and the group centroids with respect to each function provided in Table 5.17. Loadings with asterisks indicate on which function each self-concept variable has the highest loading (e.g, extraversion, excitement, fun and enjoyment, and physical vanity for Function 1 versus self-esteem for Function 2). Inspection of Table 5.17 can identify which clusters each function discriminates by examining cluster centroids and seeing where differences lie. For example, with Function 1, the centroid for Cluster 1 is -0.395, for Cluster 2 it is -0.285, and for Cluster 3 it is 0.638. From this, it can be concluded that the primary source of differences for this function is between Clusters 1 and 2 versus Cluster 3. A similar approach can be used for describing Cluster 4 and 5 versus Cluster 3.

Table 5.17 Values of Discriminant Functions at Cluster Centroids

Cluster Number of Case	Function			
	1	2	3	4
Cluster1	-0.395	-0.077	0.040	0.080
Cluster2	-0.285	-0.011	0.162	-0.124
Cluster3	0.638	-0.344	0.315	0.044
Cluster4	0.063	-0.168	-0.156	-0.024
Cluster5	0.115	0.171	0.004	0.009

Unstandardized canonical discriminant functions evaluated at group means

To summarize, using MANOVA, the present study has shown that the five clusters are different with respect to a linear combination of the nine self-concept variables. Results obtained from multivariate discriminant analysis further pinpoint which combination of self-concept variables are the most efficient in discriminating among the clusters. Based on the four criteria discussed earlier, extraversion, openness to experience, excitement, fun and enjoyment, physical vanity, achievement vanity, and self-fulfillment were found to have the most discriminating power in distinguishing among the five clusters.

To illustrate the relative merit or contribution of all nine self-concept variables, the following example shows how these variables add to the discriminating power of an already substantive discriminant model. Specifically, the nine variables are added to a model that contained past enlargement behavior as its only discriminating variable. The associated Wilks' *Lambda*, variance explained, and percentage correctly classified are computed and compared, by order of the respective variables entering the model, to assess the nine variables' contribution to the overall fitness of the full discriminant model. As shown in Table 5.18, the overall fit of the discriminant model greatly improved as indicated by all indices. Wilks' *Lambda* decreases from 0.959 to 0.872, an improvement of 9 percent. Variance explained increases from 4.1 percent to 7.6 percent, an improvement of 85 percent. The hit ratio rises from 16.4 percent to 30.1 percent, a 90 percent improvement from the previous hit ratio.

Table 5.18 Summary of Sequential Discriminant Analysis

Variables Entered	Wilks' λ	Δλ	Variance Explained (I^2)	ΔI^2	Hit Ratio	Δ%	Cross-Validated
Past Enlargement Experience	0.959		4.1%		16.4%		16.4%
Nine Self-Concept Variables	0.872	0.087	7.6%	3.5%	30.1%	14.7%	24.4%

5.7.3 Description of Derived Cluster Solution

As all issues concerning identification of the five clusters have been addressed, the next step proceeds to describe the clusters in detail. Descriptions of clusters focus on presenting individual variables that have statistical and theoretical significance. Such descriptions are accomplished by identifying variables with substantive loadings and understanding what the differences between cluster means on each variable indicate.

Table 5.19 summarizes means of key variables employed to describe the clusters. The two clustering variables, ideal-actual discrepancy and please oneself, along with the multiplicative composite variable between them, *IADPO*, qualify as descriptors in their own rights. From discriminant analysis results, extraversion, openness to experience, excitement, fun and enjoyment, physical vanity, achievement vanity, and self-fulfillment qualify as substantive discriminators for the five-cluster solution. Other variables considered in Table 5.19 include percentage of past enlargement experience and percentage and number of respondents for each cluster. Based on these variables, the five clusters are described as "status quos", "passives", "keen pursuers", "active searchers", and "prospects" as follows.

Table 5.19 Key Variables Employed to Describe Clusters

Key Variables	Cluster				
	1	2	3	4	5
Ideal-Actual Discrepancy[a,e] (*IAD*)	1.22	1.31	3.12	2.21	1.29
Please Oneself[b,e] (*PO*)	3.48	1.62	6.00	4.64	5.74
IADPO[c]	4.25	2.12	18.72	10.25	7.40
Self-Esteem[d,e]	4.42	4.47	4.51	4.39	4.63
Extraversion[d,e]	3.89	3.88	4.15	4.08	4.11
Openness to Experience[d,e]	4.35	4.28	4.65	4.43	4.36
Excitement[d,e]	4.33	4.32	4.51	4.46	4.53
Fun and Enjoyment[d,e]	4.47	4.51	4.80	4.74	4.78
Physical Vanity[d,e]	3.92	4.04	4.81	4.27	4.27
Achievement Vanity[d,e]	4.22	4.21	4.56	4.28	4.28
Being Well-Respected[d,e]	4.11	4.21	4.53	4.21	4.16
Self-Fulfillment[d,e]	4.05	4.16	4.42	4.13	4.26
Percentage of Past Enlargement Experience	15.1	16.4	46.4	35.2	20.2
Percentage of Respondents	16.8	10.7	5.1	24.8	42.6
Number of Respondents	86	55	26	127	218

[a]Based on raw data scores. 1 signifies just right; 2 one cup size larger; 3 two cup size larger; and 4 three cup size larger.
[b]Based on raw data scores. 1 signifies not at all important; 2 quite unimportant; 3 somewhat unimportant; 4 somewhat important; 5 quite important; and, 6 strongly important.
[c]Expressed as the product between *IAD* and *PO*.
[d]Based on raw data scores. 1 signifies strongly disagree; 2 disagree; 3 somewhat disagree; 4 somewhat agree; 5 agree; and, 6 strongly agree.
[e]Expressed as mean values for each cluster.

Status Quos (Cluster 1). Status quos comprise women who are satisfied with their current breast sizes and are neutral in terms of the extent to which the decision to enlarge breasts is based on a desire to please oneself (self-oriented motivation). Status quos have the second lowest *IADPO* score, followed only by the passives. Compared to keen pursuers, active searchers, and prospects, status quos join the passives in expressing less extraversion, excitement, achievement vanity, and significantly less concern about physical appearance. They join prospects in expressing a moderate degree of openness to experience. They are least oriented toward fun and enjoyment and self-fulfillment especially when compared with prospects. Made up of approximately 17 percent of the sample, the cluster has the lowest proportion of subjects with past enlargement experience.

Passives (Cluster 2). Passives are typified by women who are satisfied with their current breast sizes and present almost no self-oriented motivation toward breast enlargement. Passives have the lowest *IADPO* score; they are among the lowest scores on ideal-actual

discrepancy, and improvement of breast sizes for their own purposes is almost unimportant for them. Compared to keen pursuers, active searchers, and prospects, they join the status quos in expressing less extraversion, meaning that they are less comfortable to be surrounded by people; less excitement, meaning that they are less willing to do things out of the ordinary; and less achievement vanity, meaning that they have less concern for personal achievements. Passives have the lowest scores on openness to experience. They join status quos in having low scores on fun and enjoyment, concern for physical appearance, and self-fulfillment. Made up of approximately 10 percent of the total sample, the cluster has one of the lowest proportions of subjects with real past enlargement experience.

Keen Pursuers (Cluster 3). Keen pursuers are best characterized by considerable concern over ideal-actual discrepancy and strong self-oriented motivations compared to any other cluster. Keen pursuers have the highest *IADPO* score; they have the largest ideal-actual discrepancy and improvement of breast sizes for their own purposes is strongly important for all of them. Keen pursuers express a higher degree of extraversion and openness to experience than any other cluster, meaning that they are likely to be comfortable when surrounded by people and to be open to new experiences and ideas. Joining active searchers and prospects, they are among the highest in terms of value toward excitement. Keen pursuers join active searchers and prospects in sharing values toward fun and enjoyment in life. They perceive good looks and physical appearance (physical vanity) as more important than any other clusters. Keen pursuers are the most concerned about personal achievements compared to any other cluster. They always are looking for more challenges in work, relationships, and leisure, reflecting their orientation toward self-fulfillment. Despite being the smallest cluster with only five percent of the total sample, keen pursuers have the highest proportion of respondents experiencing breast enlargement activities in the past.

Active Searchers (Cluster 4). Active searchers are best characterized by active search for physical attractiveness. Active searchers have the second highest *IADPO* score; they have moderate concern over ideal-actual discrepancy and moderate self-oriented motivations. Active searchers join prospects in sharing extraversion and achievement vanity. They have among the highest scores in terms of openness to experience. Active searchers join keen pursuers and prospects in sharing a similar degree of value toward excitement. They pair with prospects in terms of fun and enjoyment. Active searchers join prospects in presenting concern about physical vanity but differ in terms of physical vanity, especially from status quos and keen pursuers. They join the passives in sharing a relatively low degree of self-fulfillment. Made up of approximately 25 percent of the total sample, the cluster has among the highest proportions of subjects with past enlargement experience, second only to keen pursuers.

Prospects (Cluster 5). Prospects are characterized by satisfaction toward their current sizes but have great self-oriented motivations. Prospects have the third highest *IADPO* score; although they do not see any ideal-actual discrepancy on their own breast sizes, they perceive breast enlargement as a way to satisfy themselves as strongly important. Prospects have the highest scores on self-esteem. They join active searchers in sharing extraversion and achievement vanity. They join status quos in expressing a moderate degree of openness to experience. Prospects join keen pursuers and active searchers in sharing excitement. Again, they pair with active searchers in terms of fun and enjoyment, making them significantly different from the status quos. Prospects join active searchers in presenting concern about physical vanity but they significantly differ from status quos and keen pursuers on this regard. They have among the highest scores in terms of self-fulfillment. Made up of approximately

40 percent of the total sample, prospects are the biggest cluster. The cluster has a moderate proportion of subjects with real past enlargement experience.

To summarize, this section started by showing that the five clusters are different with respect to a linear combination of nine self-concept variables. Results from MANOVA and discriminant analysis reveal that all nine self-concept variables as a combination play a role in discriminating among the five clusters. In addition, results from discriminant analysis further reveal that substantive contributions come from extraversion, openness to experience, excitement, fun and enjoyment, physical vanity, achievement vanity, and self-fulfillment. An example was given as to how the nine self-concept variables add to the discriminating power of an already powerful discriminant model. The section ended by describing the five clusters using results from a series of analyses conducted in the section. These five clusters are labeled as status quos, passives, keen pursuers, active searchers, and prospects.

Attention now moves to two-group analyses. As discussed in Chapter 3, two-group analyses of interest in this study are: (a) the most extreme cluster versus the other four clusters as a combined group; and, (b) the two most extreme clusters. In both analyses, attention will focus on significance of group differences on the nine self-concept variables.

5.8 Two-Group Analyses

Aims of this section are twofold. First, this section seeks to examine discriminating power that the entire set of self-concept variables has in differentiating between two clusters in two different cases. Second, it seeks to determine which self-concept variables are substantive in isolating one cluster from the other in those cases. In the first case, keen pursuers, the cluster with the largest *IADPO* score, are compared with all other clusters combined. In the second

case, keen pursuers are compared with passives, the cluster with the smallest *IADPO* score. Specifically, 18 *a priori* directional tests of hypotheses are conducted, as described in Hypotheses 1a to 9a and Hypotheses 1b to 9b in Chapter 3.

5.8.1 Keen Pursuers versus All Other Clusters Combined

The aim of this analysis is to compare and contrast keen pursuers with the combination of all other clusters with regard to the nine self-concept variables of interest. This analysis evaluates the contribution of the nine self-concept variables as predictors to separate keen pursuers from all other clusters. It also determines which self-concept variables are important for isolating keen pursers from the rest.

MANOVA was conducted to begin the analysis with the following results: Wilks' *Lambda* = 0.97; $F = 1.86$; $p < 0.055$. It is important to note that *p* value indicated in MANOVA reflects a multivariate two-tailed test. However, given that the all hypotheses are directional and entail one-tailed tests and that a small sample size for keen pursuers is involved ($n = 26$), it is appropriate to apply a 90 percent confidence level ($p < 0.100$) (Lind, Marchal, and Wathen 2005, p. 280; Rosnow and Rosenthal 1989). Following this procedure, it can be concluded that the present case exhibits significant differences ($p < 0.055$) on the nine self-concept variables.

To identify which self-concept variables are responsible for these differences involved discriminant analysis. As a preparatory step, it is important to ensure that the data meet assumptions required by discriminant analysis: freedom from multicollinearity, linear relationships among all pairs of predictors, and variance/covariance equality. Earlier results in Subsection 5.2.4 demonstrated that the data are free from multicollinearity. Linearity is

not a concern and tests conducted earlier in Subsection 5.7.1 reinforced irrelevance of the linearity assumption to this study. As shown in Table 5.14, maximum absolute values for skewness and kurtosis for the data are 0.52 and 0.45, lower than the general rule of thumb of 0.75 and 1.50. To examine if the variance/covariance equality assumption is violated, Box's *M* test statistic is evaluated to test the null hypothesis of the equality of variance/covariance matrices across the five groups. The following statistics are obtained: Box's $M = 38.321$; approx. $F = 0.738$; $df = 45,6116$; $p < 0.902$. Since the null hypothesis cannot be rejected, the assumption of equal variance/covariance is met.

Table 5.20 provides overall results for the two-group discriminant analysis. As expected, the discriminant function has the same level of significance as found using MANOVA ($p < 0.055$). A canonical correlation of 0.180 is obtained. Squaring the canonical correlation yields 3.2 percent of variance explained. Confirming results from MANOVA conducted above, results from discriminant analysis indicate that keen pursuers exhibit significant differences from all other clusters combined on the set of nine self-concept variables.

Table 5.20 Summary of Two-Cluster (Keen Pursuers versus the Other Clusters) Discriminant Analysis Results

Canonical Discriminant Functions

Function	Eigenvalue	Percent of Variance	Canonical Correlation	Percent of Total Variance Explained
1	0.033	100.0	0.180	3.2

Wilks' *Lambda*

Test of Function	Wilks' *Lambda*	Chi-square	*df*	*p*
1	0.968	16.608	9	0.055

To pinpoint which combination of self-concept variables are the most efficient in discriminating between the two clusters, Table 5.21 contains structure loadings for the discriminant function, along with simple correlations between each self-concept variable and

the discriminant function. Positive signs of the structure loadings indicate that as the self-concept variables increase, individuals are more likely to be classified as members of the keen pursuers group. Also presented in Table 5.21 are rankings of the self-concept variables in terms of their structure loadings. With the exception of self-esteem, all structure loadings represent significant correlations with the discriminant function ($p < 0.05$, one-tail). In addition, mean values presented in Table 5.21 indicate that keen pursuers have larger mean values than the combination of all other clusters across all self-concept variables, the only exception being self-esteem. Based on rules of thumb discussed in the previous section, openness to experience, physical vanity, achievement vanity, being well-respected, and self-fulfillment were found to have the most discriminating power that best distinguishes keen pursuers from all other clusters.

Table 5.21 Summary of Interpretive Measures for Two-Cluster Discriminant Analysis (Keen Pursuers versus All Other Clusters Combined)

Independent Variables	Mean Values		Structure Loadings		
	Keen Pursuers	All Other Clusters Combined	Value	Rank	Simple Correlation
Self-Esteem	4.51	4.51	-0.003	9	-0.003
Extraversion	4.15	4.04	0.175	6	0.178**
Openness to Experience	4.65	4.37	0.457	3	0.463**
Excitement	4.51	4.45	0.095	8	0.097*
Fun and Enjoyment	4.76	4.69	0.096	7	0.098*
Physical Vanity	4.85	4.19	0.849	1	0.853**
Achievement Vanity	4.56	4.26	0.432	4	0.438**
Being Well-Respected	4.53	4.17	0.460	2	0.466**
Self-Fulfillment	4.42	4.18	0.340	5	0.345**

*Significant at $p < 0.05$, one-tail.
**Significant at $p < 0.01$, one-tail.

Cluster centroids can be used to interpret the discriminant function results from a global perspective. Cluster centroids represent the mean discriminant function score for each group. Results reveal that the cluster centroid for keen pursuers is 0.789, whereas the cluster centroid for all other clusters combined is -0.042.

To summarize, using MANOVA, this subsection has shown that keen pursuers and the combination of all other clusters are different with respect to a linear combination of the nine self-concept variables. Results obtained from multivariate discriminant analysis indicate that all self-concept variables except self-esteem are significant predictors of group membership for keen pursuers and the combination of all other clusters (see Table 5.21). Therefore, with the exception of Hypothesis 7a, Hypotheses 1a to 9a are supported. In addition, multivariate discriminant analysis further pinpoints which combination of self-concept variables are the most effective in discriminating between the clusters. Based on criteria earlier discussed in Section 5.7.2, physical vanity, being well-respected, openness to experience, achievement vanity, and self-fulfillment, in order of their structure loadings, were found to have the most discriminating power that best distinguishes between keen pursuers and the combination of all other clusters.

5.8.2 Keen Pursuers versus Passives

This analysis aims to compare and contrast keen pursuers with passives, the two extreme clusters, with regard to the nine self-concept variables of interest. The analysis evaluates the contribution of self-concept variables as predictors to separate keen pursuers from passives.

MANOVA was conducted with the following results: Wilks' *Lambda* = 0.83; $F = 1.59$; $p < 0.134$. Again, it is important to note that the *p*-value indicated in MANOVA reflects a multivariate two-tailed test. However, given that directional hypotheses entail one-tailed tests, and that small sample sizes are involved ($n = 26$ for keen pursuers and $n = 55$ for passives), it is appropriate to apply a 90 percent confidence level (Lind, Marchal, and Wathen 2005, p. 280; Rosnow and Rosenthal 1989). Following this procedure, it can be concluded

that the two extreme clusters, keen pursuers and passives, exhibit significant differences on the nine self-concept variables.

Next, discriminant analysis was performed to examine discriminating power that individual self-concept variables have in differentiating between keen pursuers and passives. As a preparatory step, it is important to ensure that the data meet assumptions required by discriminant analysis: freedom from multicollinearity, linear relationships among all pairs of predictors, and variance/covariance equality. Earlier results in Subsection 5.2.4 demonstrated that the data are free from multicollinearity. Linearity is not a concern and tests conducted earlier in Subsection 5.7.1 reinforced irrelevance of the linearity assumption to this study. Further, there is little chance of any linear relationship being significant given the small sample size for the keen pursuers and passives clusters. As shown in Table 5.22, maximum absolute values for skewness and kurtosis for the data are 0.57 and 0.84, lower than the general rule of thumb of 0.75 and 1.5. To examine if the variance/covariance equality assumption is violated, Box's *M* test statistic is evaluated to test the null hypothesis of the equality of variance/covariance matrices across the five groups. The following statistics are obtained: Box's $M = 51.448$; approx. $F = 0.973$; $df = 45,8459$; $p < 0.524$. Since the null hypothesis cannot be rejected, the assumption of equal variance/covariance is met.

Table 5.22 Analysis of Skewness and Kurtosis (*n* = 81)

	Mean	Standard Deviation	Skewness	Standard Error	Kurtosis	Standard Error
Self-Esteem	4.48	0.91	-0.57	0.27	-0.18	0.53
Extraversion	3.96	0.85	-0.06	0.27	-0.13	0.53
Openness to Experience	4.40	0.71	-0.26	0.27	-0.09	0.53
Excitement	4.38	0.89	-0.13	0.27	-0.74	0.53
Fun and Enjoyment	4.60	0.96	-0.35	0.27	-0.35	0.53
Physical Vanity	4.30	1.07	0.10	0.27	-0.84	0.53
Achievement Vanity	4.33	0.91	-0.03	0.27	-0.76	0.53
Being Well-Respected	4.31	1.06	-0.48	0.27	0.11	0.53
Self-Fulfillment	4.25	0.71	0.27	0.27	-0.28	0.53

Table 5.23 provides overall results for the two-cluster discriminant analysis. As expected, the discriminant function is significant at the same *p*-value found for the MANOVA analysis. A canonical correlation of 0.410 is obtained which yields 16.8 percent of variance explained. Confirming results from MANOVA conducted above, results from discriminant analysis indicate that keen pursuers exhibit significant differences from passives on the set of the nine self-concept variables.

Table 5.23 Summary of Two-Cluster (Keen Pursuers versus Passives) Discriminant Analysis Results

Canonical Discriminant Functions

Function	Eigenvalue	Percent of Variance	Canonical Correlation	Percent of Total Variance Explained
1	0.202	100.0	0.410	16.8

Wilks' *Lambda*

Test of Function	Wilks' *Lambda*	Chi-square	*df*	*p*
1	0.832	13.678	9	0.134

To identify which self-concept variables are the most efficient in discriminating between the two extreme clusters, Table 5.24 contains structure loadings for the discriminant function, along with simple correlations between each self-concept variable and the discriminant function. Positive signs of the structure loadings indicate that as the self-concept variables increase, individuals are more likely to be classified as members of the keen pursuers group. Also presented in Table 5.24 are rankings of the self-concept variables in terms of their structure loadings. With the exception of self-esteem, all structure loadings represent significant correlations with the discriminant function ($p < 0.05$, one-tail). In addition, mean values presented in Table 5.24 indicate that keen pursuers have larger mean values than passives across all self-concept variables. Based on rules of thumb discussed in the previous section, extraversion, openness to experience, physical vanity, achievement vanity, being

well-respected, and self-fulfillment have the most discriminating power in distinguishing keen pursuers from passives.

Table 5.24 Summary of Interpretive Measures for Two-Cluster Discriminant Analysis (Keen Pursuers versus Passives)

Independent Variables	Mean Values		Structure Loadings		
	Keen Pursuers	Passives	Value	Rank	Simple Correlation
Self-Esteem	4.51	4.47	0.048	9	0.052
Extraversion	4.15	3.88	0.337	5	0.365**
Openness to Experience	4.65	4.28	0.569	2	0.604**
Excitement	4.51	4.32	0.233	8	0.254*
Fun and Enjoyment	4.76	4.52	0.259	7	0.282*
Physical Vanity	4.85	4.04	0.850	1	0.870**
Achievement Vanity	4.56	4.21	0.412	3	0.445**
Being Well-Respected	4.53	4.21	0.313	6	0.340**
Self-Fulfillment	4.42	4.16	0.387	4	0.418**

*Significant at $p < 0.05$, one-tail.
**Significant at $p < 0.01$, one-tail.

The cluster centroid for keen pursuers is 0.645 and the cluster centroid for passives is -0.305.

To summarize, this subsection has shown that keen pursuers and passives are different with respect to a linear combination of the nine self-concept variables. Results obtained from multivariate discriminant analysis point out that all self-concept variables except self-esteem are significant predictors of group membership for keen pursuers and passives (see Table 5.24). Therefore, with the exception of Hypothesis 7b, Hypotheses 1b to 9b are supported. In addition, multivariate discriminant analysis further pinpoints which combination of self-concept variables are the most efficient in discriminating between the clusters. Based on criteria earlier discussed in Section 5.7.2, physical vanity, openness to experience, achievement vanity, self-fulfillment, extraversion, and being well-respected, in an order of their structure loadings, were found to have the most discriminating power that best distinguishes between keen pursuers and passives.

5.9 Summary of Hypotheses Testing Results

This section summarizes statements of hypothesis of the present study and their test results. As has been described throughout the chapter, the derived five-cluster solution obtained from *k*-means clustering has been shown to be different with respect to a combination of the nine self-concept variables. In addition, both cases of two-group comparisons also demonstrated distinctions with regard to a combination of the nine self-concept variables. A series of test results obtained from MANOVAs and discriminant analyses have confirmed these distinctions, as summarized in Table 5.25.

Hypotheses 1 to 9 seek to identify which variables, among the nine self-concept variables, are significant predictors of group membership in the two cases of two-group comparisons—keen pursuers versus the combination of all other clusters (Hypotheses 1a to 9a) and keen pursuers versus passives (Hypotheses 1b to 9b). Results of the 18 tests indicate that all self-concept variables are good predictors with the exception of self-esteem.

Table 5.25 Summary of Hypotheses and Test Results

Hypothesis	Content of Hypothesis	Hypothesis Supported
H1a, H1b	Fun and enjoyment is positively associated with discriminant function *Z* scores that classify female university students into segments based on ideal-actual discrepancy and self-oriented motivation.	Yes
H2a, H2b	Excitement is positively associated with discriminant function *Z* scores that classify female university students into segments based on ideal-actual discrepancy and self-oriented motivation.	Yes
H3a, H3b	Being well-respected is positively associated with discriminant function *Z* scores that classify female university students into segments based on ideal-actual discrepancy and self-oriented motivation.	Yes
H4a, H4b	Self-fulfillment is positively associated with discriminant function *Z* scores that classify female university students into segments based on ideal-actual discrepancy and self-oriented motivation.	Yes
H5a, H5b	Extraversion is positively associated with discriminant function *Z* scores that classify female university students into segments based on ideal-actual discrepancy and self-oriented motivation.	Yes
H6a, H6b	Openness to Experience is positively associated with discriminant function *Z* scores that classify female university students into segments based on ideal-actual discrepancy and self-oriented motivation.	Yes
H7a, H7b	Self-Esteem is negatively associated with discriminant function *Z* scores that classify female university students into segments based on ideal-actual discrepancy and self-oriented motivation.	No
H8a, H8b	Physical vanity is positively associated with discriminant function *Z* scores that classify female university students into segments based on ideal-actual discrepancy and self-oriented motivation.	Yes
H9a, H9b	Achievement vanity is positively associated with discriminant function *Z* scores that classify female university students into segments based on ideal-actual discrepancy and self-oriented motivation.	Yes

5.10 Chapter Summary

Chapter 5 aims to test the validity of the derived cluster solution through a series of hypotheses as stated in Chapter 3. A series of tests were conducted by comparing and contrasting two cases of two-group comparisons with respect to the nine self-concept variables. Test results support hypotheses for all self-concept variables except self-esteem. The nine self-concept variables discriminate well either in the case of five-cluster solution or in both cases of two-group analyses. Discussions of test results that incorporate the theoretical perspective of self-concept will be conducted in Chapter 6.