The results from all the relevant analyses provided support for the validation of the proposed model. Confirmatory factor analysis were used to assess the unidimensionality, reliability, and validity of the scale (Steenkamp & Van Trijp, 1991:283). The overall model goodness-of-fit results and the measurement model assessments lent substantial support for confirmation of the proposed four-factor model.

7.5 HYPOTHESES TESTING

The process of hypotheses testing was conducted as follows: first, statistical hypotheses were determined by formulating null and alternative hypotheses (as set out in Chapter 5). The next step was to specify the circumstances under which H_0 would or could not be rejected, by choosing a level of significance. A significance level is a critical probability in choosing between the null hypothesis and the alternative hypothesis. A five per cent significance level (α=0.05) was set for all hypotheses. Thereafter an appropriate statistical technique with a corresponding test statistic was chosen. Finally, the values of the test statistics were calculated, the test results interpreted and a decision was made to reject or not reject the null hypotheses. All the significant results are indicated in bold print.

All the hypotheses that were tested with the same statistical technique are discussed in the same section. This means that the hypotheses do not necessarily follow a chronological order. The following section provides detailed results for the hypotheses testing based on the above-mentioned principles.

7.5.1 Testing hypotheses using chi-square tests

Hypotheses 2b, 3b, 3c and 6 were tested by means of chi-square (χ^2) tests. In H_{2b} and H_{3b}, two groups were compared on a variable measured on a nominal scale and were therefore tested with the two-sample chi-square test for independency (Sections 7.4.1.1 and 7.4.1.2). H_{3c} was tested with the k-sample chi-square test for independency (an extension of the two-sample chi-square test) because comparisons were made between

more than two groups (Section 7.4.1.3). The null hypotheses tested by the chi-square tests for independence was that there is no difference between the groups in respect of the relative frequency with which group members belonged to the various categories of the variables of interest. The reason for focusing on relative rather than absolute frequencies was that the groups had unequal sample sizes and, therefore, the calculation of expected frequencies needed to take this into account, as suggested by Diamantopoulos and Schlegelmilch (1997:175). With a 2x2 table (as is the case with H_{2b} and H_{3b}), it is often recommended that Yates's correction for continuity be applied to obtain a modified chi-square statistic (Diamantopoulos & Schlegelmilch, 1997:177). This is designed to correct or compensate for what some researchers regard as an overestimate of the chi-square value when used with a 2x2 table. Yates' correction for continuity is reported in both H_{2b} and H_{3b} (see Sections 7.4.1.1 and 7.4.1.2). H_6 was tested by means of the chi-square goodness-of-fit test. Here the 'goodness of fit' of the observed distribution with the expected distribution was tested. The null hypothesis under the chi-square goodness-of-fit test was that the observed frequencies were equal to the theoretical frequencies. With the testing of H_6, the expected frequencies were all specified to be the same (see discussion in Section 7.4.1.4).

Two assumptions underlie chi-square tests (Diamantopoulos & Schlegelmilch, 1997:180; Keller & Warrack, 2000:557). First, the observations must originate from a random sample and the scores associated with the observation must be independent from each other. The second assumption is that the sample size must be relatively large so that a minimum expected cell frequency can be obtained. There is no simple answer to the question of what sample size is large enough, but many researchers suggest that at least 80 per cent of the cells should have an expected frequency greater than or equal to 5, and that no cell should have an expected frequency smaller than 1 (Diamantopoulos & Schlegelmilch, 1997:180; Green, Salkind & Akey, 1999:346; Pallant, 2001:257). The data did not violate any of the stated assumptions. The results for the first four hypotheses are presented below.

7.5.1.1 *H_{2b}: There is a dependency between being a victim of invasion of privacy and gender*

Hypothesis 2b attempted to find support that there are differences between gender groups in terms of their privacy invasion experiences. Results from previous international studies had indicated that men seemed to be more likely than women to report being a victim of privacy invasion (refer to Chapter 5, Section 5.4.2). The result of the two-sample chi-square for the data in this study is illustrated in Table 7.24.

Table 7.24 Difference between gender and victims of privacy invasion

VICTIMS OF INVASION GENDER		Victim of privacy invasion	Not a victim of privacy invasion	ROW TOTAL	
Male	Frequency Row %	89 38	148 62	237	
Female	Frequency Row %	107 27	283 73	390	
COLUMN TOTAL		196	431	627	
p-value					**0.0104**

The specified significance level of 5 per cent resulted in a corrected p-value of 0.0104 (Yates's correction for continuity), **leading to a rejection of the null hypothesis**. There was support for H_{2b}, indicating that there is a difference between the number of privacy invasions reported between men and women. To establish what percentage of males versus females reported being a victim of privacy invasion, the row percentages needed to be interpreted. From Table 7.24, it can be deduced that 38 per cent of males were victims of privacy invasion, while 62 per cent were not victims. For females, 27 per cent reported being a victim, versus 73 per cent that were not victims. This indicates that males are more likely to perceive themselves as victims of privacy invasion (38% versus 27%). This result corresponds with findings from previous studies which indicated that **males seem to be more likely than females to perceive themselves as victims of privacy invasion**.

7.5.1.2 *H_{3b}: There is a dependency between the level of awareness of name removal procedures and age*

There is sufficient evidence in previous literature to suggest that there are differences between lower and higher age groups in terms of their awareness of name removal procedures (refer to Chapter 5, Section 5.4.3). Hypothesis 3b was formulated to seek support for age differences of consumers in terms of their knowledge of name removal procedures. Question 61 in the questionnaire (see Appendix 1) requested respondents to provide their year of birth as a 4-digit number, for example 1969. Answers to this question, based on the frequencies, were categorised into two distinct age groups, namely a younger age group (18-39 years) and an older age group (40+ years). The result of the chi-square test comparing younger and older consumers in terms of their awareness of name removal procedures is presented in Table 7.25.

Table 7.25 Difference between age groups and awareness of name removal procedures

AWARENESS OF NAME REMOVAL / AGE GROUPS		Aware of name removal procedures	Not aware of name removal procedures	ROW TOTAL	
18-39 years	Frequency	64	215	279	
	Row %	23	77		
40+ years	Frequency	80	266	346	
	Row %	23	77		
COLUMN TOTAL		144	481	625	
p-value					1.0000

Since the corrected p-value (Yates's correction for continuity) was larger than the specified significance level of 5 per cent, the null hypothesis could not be rejected. There was thus **not enough empirical support for H_{3b} to suggest that the level of awareness of name removal procedures is dependent on age.** For both groups, a total of 77 per cent of the consumers were unaware that they can remove their names from some of the major contact lists in the country as a means to protect their privacy.

7.5.1.3 *H_{3c}: There is a dependency between the awareness of name removal procedures and levels of education*

Hypothesis 3c was formulated on the basis of evidence from previous studies which had found that consumers who are aware of name removal procedures are often better educated than those who are not aware of name removal procedures (refer to Chapter 5, Section 5.4.3). Answers to question 63 in the questionnaire (see Appendix 1) were regrouped to form three different educational groups, namely low, medium and high levels of education. Table 7.26 presents the result of the chi-square test indicating the differences between the groups with low, medium and high educational levels in terms of their awareness of name removal procedures.

Table 7.26 Difference between education levels and awareness of name removal procedures

AWARENESS OF NAME REMOVAL PRODECURES / EDUCATION LEVELS		Aware of name removal procedures	Not aware of name removal procedures	ROW TOTAL	
Low educational level (Up to Grade 10)	Frequency	35	119	154	
	Row %	23	77		
Medium educational level (Up to Grade 12)	Frequency	54	179	233	
	Row %	23	77		
High educational level (Post Grade 12 qualification)	Frequency	56	182	238	
	Row %	24	76		
COLUMN TOTAL		145	480	625	
p-value					0.9832

The p-value indicates that the null hypothesis cannot be rejected as the results provide no support for H_{3c}. From this non-significant p-value, one can conclude that **there is no relationship between educational levels and awareness of name removal procedures**. This result differs from that of previous studies where more highly educated consumers were more aware of name removal procedures than less educated consumers.

7.5.1.4 H_6: The proportion of South African consumers is not equally represented in the different privacy segments

Based on the Privacy Segmentation Index designed by Westin and Louis Harris & Associates (Harris Interactive, 2002b:20), respondents were categorised into one of three segments, depending on their degree of agreement of disagreement with three questions (refer to Chapter 5, Section 5.4.6). Appendix 1 includes Questions 46 to 48 used in this analysis. Respondents who 'strongly agreed' or 'slightly agreed' with Question 46, and 'strongly disagreed' or 'slightly disagreed' with Questions 47 and 48, were labelled 'Privacy Fundamentalists'. Respondents who 'strongly disagreed' or 'slightly disagreed' with Question 46, and 'strongly agreed' or 'slightly agreed' with Questions 47 and 48 were labelled 'Privacy Unconcerned'. The remaining options were grouped to form the third segment and were labelled 'Privacy Pragmatists'. The frequencies and percentages for the different privacy segmentation groups are shown in Table 7.27.

Table 7.27 Frequencies and percentages for the different privacy segments

PRIVACY SEGMENTS	n	%
Privacy Fundamentalists	191	30.46
Privacy Unconcerned	69	11.00
Privacy Pragmatists	367	58.53
Total	627	100.00

From Table 7.27 it is clear that more than 50 per cent of the respondents belong to the Privacy Pragmatist segment. In order to empirically test H_6 for the proportions of the three segments, a chi-square goodness-of-fit test was conducted. The results are shown in Table7.28.

Table 7.28 Chi-square test results of privacy segments

PRIVACY SEGMENTS	Privacy Fundamentalists	Privacy Unconcerned	Privacy Pragmatists	Total	
Observed frequencies	191	69	367	627	
Expected frequencies	209	209	209	627	
p-value					0.0000

The chi-square test resulted in a p-value of 0.0000, indicating that there are significant differences between the segments. The null hypothesis under the chi-square goodness-of-fit test is that the observed frequencies are equal to the theoretical frequencies, therefore the expected frequencies were all specified to be equal (refer to Table 7.28). When the null hypothesis is true, the observed and expected frequencies should be similar, in which case the test statistic would be small. The chi-square result provides sufficient evidence to infer that the proportions are not the same. **This chi-square result, together with the fact that 59 per cent of South African consumers fall within the Privacy Pragmatists segment, 30 per cent within the Privacy Fundamentalists segment, and 11 per cent within the Privacy Unconcerned segment, provide enough evidence that the null hypothesis can be rejected, leading to support for H_6.**

The remainder of the hypotheses formulated in Chapter 5 were all tested by means of the same statistical test (MANOVA) and are discussed in the next section.

7.5.2 Testing hypotheses using MANOVA

The final phase in the data analysis was to test Hypotheses 1, 2a, 3a, 4, 5 and 7a to 7f. Multiple analyses of variance (MANOVA) were performed to address the remainder of the research hypotheses. All these hypotheses related to overall privacy concerns, and instead of using simple ANOVAs on each identified factor (dependent variable), MANOVA was used to take into account the pattern of covariation among all four of the factors, identified with the exploratory factor analysis, at the same time. The objective of

MANOVA is to test for differences in the mean values of several dependent variables across groups. This enables researchers to make inferences about whether the observed differences in the sample means across two or more groups are significant (Lattin *et al.*, 2003:389, 409).

For each of the hypotheses, a MANOVA was conducted to assess the differences between the groups in terms of their overall privacy concerns (Factors 1 to 4 simultaneously). Because the multivariate test of MANOVA shows only an overall significant difference, univariate analyses and *post hoc* comparisons were performed to reveal more specific differences between groups on each of the identified factors (Factors 1 to 4 individually). To test for differences between groups for each identified hypothesis (Hypotheses 1, 2a, 3a, 4, 5 and 7a to 7f), several groups had to be formed. These include:

- protective behaviour (three subgroups);
- privacy victims (two subgroups);
- awareness of name removal procedures (two subgroups);
- Internet usage (two subgroups);
- direct shopping (two subgroups);
- age (two subgroups);
- language (three subgroups);
- education (three subgroups);
- employment (two subgroups);
- income (three subgroups); and
- gender (two subgroups).

Detail on the formation of each subgroup are addressed later in the chapter, when the testing each individual hypothesis is discussed.

Before MANOVA could be conducted, certain critical assumptions about the nature of the data needed to be addressed.

7.5.2.1 *Assumptions of MANOVA*

For the multivariate test procedures of MANOVA to be valid, three assumptions must be met: the observations must be independent, the set of dependent variables must follow a multivariate normal distribution, and the variance-covariance matrices must be equal for all treatment groups (Hair *et al.*, 1998:347).

(a) Assumption 1

The most basic violation of the MANOVA assumption occurs when there is a lack of independence among observations. The independence of the respondents was relatively ensured by the random sampling plan (refer to Section 6.5 in Chapter 6).

(b) Assumption 2

The second assumption for MANOVA to be valid concerns the normality of the dependent measures. Significance tests for MANOVA are based on the multivariate normal distribution. Multivariate normality implies that the sampling distributions of the means of the various dependent variables in each cell are normally distributed (Hair *et al.*, 1998:73). The Kolmogorov-Smirnov test was conducted to assess the normality of the dependent variables. This test calculates the level of significance for the differences from a normal distribution and, if the test statistic is significant, it indicates that there is not a normal distribution. The different Kolgomorov-Smirnov tests for the sampling distributions of the means for the various dependent variables are set out in Appendix 7. Despite the fact that all the test results (except for one) were significant (indicating non-normality), a decision was made to accommodate the normality violation for several reasons. First, violations of this assumption have little impact with larger sample sizes, as is the case in the current study. Second, the normality violation can be accommodated as long as the differences are not due to outliers (Hair *et al.*, 1998:349; Tabachnick & Fidell, 2001:329). The data were examined for outliers, and a visual examination of the data did not indicate any outliers in the data. Third, Tabachnick and Fidell (2001:329) suggested that when the smallest cell has 20 or more observations, even if there are unequal cell sizes, tests are robust to the violation of the normality

assumption. All cell sizes had more than 20 observations, with the smallest cell size containing 139 observations. All the afore-mentioned aspects provided support for the decision to continue with the MANOVA despite the violations of the normality assumption.

(c) Assumption 3

The last critical assumption concerns the homogeneity of the variance-covariance matrices across the groups. In MANOVA, the focus is the variance-covariance matrices of the dependent measures for each group. The first analysis assessed the univariate homogeneity of variance across the groups for the different hypotheses. The most commonly used test to assess homogeneity, the Levene test, was used to assess whether the variance of a single metric variable is equal across the applicable number of groups (Hair *et al.*, 1998:75). The Levene test performs an analysis of variance on the absolute deviations of values from the respective group means for each dependent variable. If the Levene test is statistically significant, then there are no homogeneous variances. The results for the different Levene tests are shown in Table 7.29.

Table 7.29 Tests for equal variances of different groups

GROUP CONCERNS	Levene's test				Box *M*
	Privacy protection (factor 1)	Information misuse (factor 2)	Solicitation (factor 3)	Government protection (factor 4)	
Protective behaviour groups	0.0000	0.0001	0.0083	**0.9748**	0.0000
Privacy victim groups	0.0066	0.0000	0.0006	**0.1791**	0.0000
Awareness of name removal	**0.5739**	0.0001	0.0000	0.0000	0.0000
Internet usage groups	**0.9711**	0.0099	0.0086	**0.9856**	0.0010
Direct shopping groups	0.0011	0.0281	0.0000	**0.1096**	0.0000
Age groups	0.0000	**0.1221**	**0.8046**	**0.9101**	0.0000
Language groups	0.0000	**0.2416**	0.0251	0.0124	0.0000
Educational groups	0.0000	0.0005	0.0013	**0.8716**	0.0000
Employment groups	**0.3319**	**0.9525**	**0.1855**	0.0197	0.0000
Income groups	0.0075	0.0075	0.0065	0.0451	0.0000
Gender groups	0.0009	**0.2828**	**0.4975**	0.0000	0.0000

As can be seen from Table 7.29, many of the Levene tests indicated unequal variances (equal variances are indicated in bold print). However, it must be noted that the Levene test is itself not necessarily very robust against violations of the homogeneity of variances assumption (StatSoft, 1995:1709).

The next step was to assess the dependent variables collectively by testing the equality all the variance-covariance matrices between the groups. The test for overall equivalence of the variance-covariance matrices is the Box *M* test. If this test is statistically significant, then the variance-covariance matrices in the different between-group cells in the design are significantly different from each other. Results of the Box *M* test for the different MANOVAs are also reported in Table 7.29 above. It is important to note that the Box *M* test is very sensitive to deviations from the normal distribution and therefore the results of this test in this study should be viewed with some scepticism, according to Tabachnick and Fidell (2001:80), who report that the Box *M* is too strict with large sample sizes. They also believe that if sample sizes are equal, significance tests can be expected to be robust and the outcome of the Box *M* can be discarded (Tabachnick & Fidell, 2001:330). Several researchers have concluded that the homogeneity of variances-covariances assumption usually does not seriously threaten the validity of the multivariate results (StatSoft, 1995:1711).

A great deal of research has assessed the robustness (or lack thereof) of ANOVA and MANOVA analyses to the violation of homogeneity of variance (Tabachnick & Fidell, 2001:80). Many believe that the formal tests of homogeneity of variance are too strict because they also assess normality. Tabachnick and Fidell (2001:80) recommend that the homogeneity of variance be assessed with F_{max} in conjunction with sample-size ratios. F_{max} is the ratio of the largest cell variance to the smallest. If sample sizes are relatively equal (within a ratio of 4 to 1 from largest cell variance to the smallest), an F_{max} as large as 10 is acceptable. In view of the fact that the F_{max} in conjunction with sample-size ratios for the data was all acceptable (refer to Appendix 8), indicating a degree of homogeneity of variance compared to the stricter Levene and Box *M* tests, a

decision was made not to transform the data. Although data transformations are recommended as a remedy for failures of normality and homogeneity, it is not universally recommended, since it limits interpretation due to the transformed scores (Tabachnick & Fidell, 2001:80). However, despite the fact that the homogeneity of variances-covariances assumption usually does not seriously threaten the validity of the multivariate results, it was decided to follow another option, namely to use the untransformed variables but to set a more stringent alpha level – using 0.025 instead of 0.05 (Tabachnick & Fidell, 2001:80).

7.5.2.2 The MANOVA process

The main purpose of using the MANOVA to test each of the hypotheses, was to assess the differences between the groups collectively rather than individually using univariate tests. Wilks' lambda was the test statistic used to assess the overall significance of the MANOVA. There is ample evidence that Wilks' lambda is the measure that is, of all the available tests, the one that is most immune to violations of the assumptions underlying MANOVA without compromising on power (Hair *et al.*, 1998:351). The larger the between-groups dispersion, the smaller the value of Wilks' lambda and the greater the implied significance. Although the multivariate tests of MANOVA do allow the rejection of the null hypotheses that the groups' means are all equal, they do not pinpoint where the significant differences lie if there are more than two groups. Therefore, where a significant Wilks' lambda result was found, it was followed by univariate analyses and *post hoc* comparisons to reveal more specific differences between groups. Where this result displayed significance, it was subjected to a one-way analysis of variance (ANOVA). In the case of more than two groups, a Scheffé *post hoc* test was applied to identify which groups displayed significant differences. An alpha level of 0.025 was specified for all the hypotheses (for the reasons set out in Section 7.4.2.1) and all significant results are indicated in bold print in the different tables.

The results of the different MANOVAs for Hypotheses 1, 2a, 3a, 4, 5 and 7a to 7f are discussed in detail below.

7.5.2.3 *H_1: There is a significant difference between consumers in terms of their protective behaviour and their privacy concerns*

Questions 49 to 53 in the questionnaire (see Appendix 1) addressed respondents' protective behaviour in terms of their information privacy. In order to make respondents' answers to these five questions more digestible, respondents were divided into one of three groups according to their 'yes' or 'no' answers to each of the five statements. If respondents answered 'no' to all five behaviour questions, they were classified as displaying 'no protective behaviour'. If respondents answered 'yes' to only one or two of the five behaviour questions, they were classified as displaying 'limited protective behaviour'. The remainder of the respondents, those who answered 'yes' to three, four or all five of the behaviour questions, were classified as a 'protective behaviour' group. Hypothesis 1 was formulated to test for significant differences between different protective behaviour groups in terms of their overall privacy concerns. Results of the MANOVA, univariate analysis and *post hoc* comparisons are set out in Table 7.30.

Table 7.30 Mean values and MANOVA results for different behaviour groups

PRIVACY CONCERNS BEHAVIOUR GROUP	Privacy protection	Information misuse	Solicitation	Government protection	F value	p-value
No protective behaviour	4.6[a]	3.5[a]	3.4[a]	4.5		
Limited protective behaviour	4.8[a]	3.7[b]	3.6[b]	4.6		
Protective behaviour	4.7[b]	4.0[ab]	3.8[a]	4.6		
Univariate analyses	**0.0000**	**0.0000**	**0.0037**	0.8412		
Wilks' lambda					6.30	**0.0000**

[a] and/or [b]: The results of the Scheffé *post hoc* tests are indicated with [a] and/or [b]. All mean values containing the same letters (for example, [a]) indicate that the groups differ significantly from one another. All mean values containing different letters (for example, an [a] or [b]) indicate that these groups do not differ significantly from one another.

The Wilks' lambda value indicates a significant difference (p=0.0000) between behaviour groups and their overall privacy concerns, providing support for H_1. The follow-up univariate analyses revealed that these differences were significant for the first three factors, namely privacy protection, information misuse and solicitation. The Scheffé *post hoc* tests revealed that the 'no protection group' differed significantly from the 'limited protection group' on Factor 1 (privacy protection). There were significant differences between both the 'no protection and limited protection groups', and the 'protective behaviour group' on Factor 2 (information misuse). Finally there were also significant differences between the 'no protective behaviour group' and the 'protective behaviour group' in terms of Factor 3 (solicitation). The high mean values for Factor 4 (government protection) indicate that **all three the behaviour groups had very strong feelings about government protection resulting in no significant differences between the groups.**

7.5.2.4 *H_{2a}: There is a significant difference between consumers who have been victims of invasions of privacy and consumers who have not been victims of invasions of privacy in terms of their privacy concerns*

Previous empirical research suggests that consumers who have been victims of privacy invasion have higher privacy concerns than consumers who have not been victims of privacy invasion (refer to Chapter 5, Section 5.4.2). H_{2a} was formulated to determine whether there would be significant differences between the victims and non-victims in terms of their privacy concerns. Results of the MANOVA as well as the subsequent univariate analyses are reported in Table 7.31.

Table 7.31 Mean values and MANOVA results for different privacy victim groups

PRIVACY CONCERNS VICTIM GROUPS	Privacy protection	Information misuse	Solicitation	Government protection	F value	p-value
Victim of privacy invasion	4.8	4.2	3.8	4.6		
Not a victim of privacy invasion	4.7	3.5	3.5	4.5		
Univariate analyses	0.1012	**0.0000**	**0.0000**	0.1829		
Wilks' lambda					17.06	**0.0000**

A total of 31 per cent of respondents indicated that they had been victims of privacy invasion (refer to Table 7.12). The MANOVA result shows significant differences between the different victim groups in terms of their overall privacy concerns. **The p-value (p=0.0000) indicated that the null hypothesis could be rejected, providing support for H_{2a}.** The follow-up univariate analyses revealed that these differences were significant for Factor 2 (information misuse) and Factor 3 (solicitation). In both cases, the victims of privacy invasion had higher mean values than the non-victims, which corresponds with the results from previous studies which indicated higher concerns among victims of privacy invasion. It is interesting to note that both groups regarded it as important to receive privacy and government protection (Factor 1 and 4) and no significant differences were identified for these two subdimensions.

7.5.2.5 *H_{3a}: There is a significant difference between consumers in terms of their level of awareness of name removal procedures and their privacy concerns*

Various studies have measured whether consumers' knowledge levels of privacy policies and practices affect their privacy concerns. Some studies have found that knowledgeable consumers are less concerned about their personal information, whereas other studies suggested that consumers are more concerned about the

collection and use of their personal information (Culnan, 1995:14; Campbell, 1997:46). In this study, consumers' knowledge was assessed by the following question: 'are you aware of any options to remove your name from records of companies?' The findings are set out in Table 7.32.

Table 7.32 Mean values and MANOVA results for the different awareness groups

PRIVACY CONCERNS AWARENESS GROUPS	Privacy protection	Information misuse	Solicitation	Government protection	F value	p-value
Aware of name removal procedures	4.8	3.5	3.3	4.3		
Not aware of name removal procedures	4.7	3.8	3.7	4.6		
Univariate analyses	0.2580	**0.0009**	**0.0027**	**0.0000**		
Wilks' lambda					12.56	**0.0000**

The Wilks' lambda value indicated a significant difference (p=0.0000) between the different awareness groups in terms of their overall privacy concerns, providing support for H_{3a}. The follow-up univariate analyses revealed that these differences were significant for all the factors, except for Factor 1 (privacy protection) where both groups had high concerns in terms of privacy protection. The significant dimensions indicated that the respondents who were **not** aware of name removal procedures were more concerned about their privacy (see higher mean values in Table 7.32).

7.5.2.6 *H_4: There is a significant difference between Internet users and Internet non-users in terms of their privacy concerns*

Questions 56 and 57 measured respondents' involvement in Internet transactions (refer to Appendix 1). If respondents answered 'no' to both questions, indicating that they were not involved in Internet transactions, they were classified as 'Internet non-users'. If

respondents answered 'yes' to one or both of the questions, indicating Internet involvement, they were classified as 'Internet users'. The mean values of the different groups and the MANOVA results are shown in Table 7.33.

Table 7.33 Mean values and MANOVA results for Internet user groups

PRIVACY CONCERNS / INTERNET USER GROUPS	Privacy protection	Information misuse	Solicitation	Government protection	F value	p-value
Internet users	4.7	4.0	3.8	4.5		
Internet non-users	4.7	3.6	3.5	4.5		
Univariate analyses	0.9837	**0.0010**	0.0344	0.9548		
Wilks' lambda					3.29	**0.0111**

The significant Wilks' lambda (p=0.0111) provided enough support for H_4, namely that Internet users and Internet non-users differ in terms of their overall privacy concerns. The subsequent univariate analyses revealed that these differences were mainly due to the information misuse dimension of privacy (Factor 2). Here the Internet users had a higher mean value than the Internet non-users, indicating that they were more concerned about the misuse of their personal information. Both Internet and Internet non-users felt that it was very important to receive privacy protection and government protection (Factor 1 and 4) as can be seen from the high alpha values and the similar mean values reported for these two subdimensions in the univariate analyses.

7.5.2.7 H_5: There is a significant difference between direct shoppers and non-direct shoppers in terms of their privacy concerns

Questions 58 to 60 in the questionnaire (see Appendix 1) related to direct purchasing – be it shopping by means of a catalogue, or placing an order telephonically. Respondents who answered 'no' to all the direct purchasing questions indicated that they had not bought anything directly during the past year, and were classified as 'non-

direct shoppers'. All the respondents who answered 'yes' to one, two or all three of the questions were classified as 'direct shoppers' because they had purchased directly before. There are no published research findings about whether direct purchasing experience and knowledge of consumers increases privacy concerns among South Africans. Results from studies in other countries have indicated that consumers who have been involved in direct purchasing tended to be more concerned about threats to their privacy (refer to Chapter 5, Section 5.4.5). The MANOVA results for H_5 are illustrated in Table 7.34.

Table 7.34 Mean values and MANOVA results for different direct shopping groups

PRIVACY CONCERNS DIRECT SHOPPING GROUPS	Privacy protection	Information misuse	Solicitation	Government protection	F value	p-value
Direct shoppers	4.8	3.6	3.3	4.5		
Non-direct shoppers	4.7	3.8	3.8	4.6		
Univariate analyses	0.0456	0.0487	**0.0000**	0.0754		
Wilks' lambda					11.52	**0.0000**

The specified significance level of 2.5 per cent resulted in a p-value of 0.0000, indicating support for H_5. This demonstrates that there is a difference between direct shoppers and non-direct shoppers in terms of their privacy concerns. To establish where the differences lie, univariate analyses were conducted on the different dependent variables (factors). This revealed that the differences between direct and non-direct purchasers were related to solicitation (Factor 3). The mean values in Table 7.34 further indicate that the non-direct shoppers were more concerned about solicitation than the direct shoppers (3.8 versus 3.3). This suggests why these respondents were classified as non-direct shoppers – they specifically do not purchase by means of catalogues, direct mail or telemarketing because they view these

purchasing forms as intrusive and do not want unsolicited communication from companies.

7.5.2.8 *H_{7a}: There is a significant difference between young and old people in terms of their privacy concerns*

Several studies on information privacy have found a relationship between age and privacy concerns (refer to Chapter 5, Section 5.4.7.1). The majority of these studies show that the level of privacy concerns increases with age. As has been mentioned in Section 7.4.1.2, respondents were divided into two age groups. The one group represented young consumers (18-39 years) and the other group represented the older consumers (40+ years). The mean values of the two age groups and the MANOVA results of the hypothesis test are shown in Table 7.35.

Table 7.35 Mean values and MANOVA results for different age groups

PRIVACY CONCERNS / AGE GROUPS	Privacy protection	Information misuse	Solicitation	Government protection	F value	p-value
18-39 years	4.7	3.6	3.5	4.5		
40+ years	4.8	3.8	3.7	4.6		
Univariate analyses	**0.0050**	**0.0157**	**0.0066**	0.3774		
Wilks' lambda					3.69	**0.0056**

The MANOVA result showed significant differences between the younger and older age groups in terms of their overall concerns. The null hypothesis was thus rejected, as there is support for H_{7a}. The follow-up univariate analyses revealed that these differences were significant for all the privacy dimensions, except for Factor 4 (government protection) where both age groups had high concerns. The higher mean values among the older age group corresponded with some of the findings of the international studies, namely that older consumers are more concerned about information privacy than younger consumers.

7.5.2.9 *H_{7b}: There is a significant difference between the main language groups in terms of their privacy concerns*

Question 62 in the questionnaire (see Appendix 1) requested respondents to indicate their home language. The question made provision for all eleven official South African languages. After an examination of the frequencies of all eleven different language groups (see Table 7.5), it was decided to reduce the eleven language groups to three main language groups to simplify the results. The first group included all English-speaking respondents, the second group all Afrikaans-speaking respondents, and the third group all the Black African language groups (refer to Table 7.5). Since the nine Black African languages accounted for 20 per cent of the sample, these languages were grouped together and labelled the Black African language group. Respondents who indicated that their home language did not belong to any one of the eleven official languages were excluded from further analysis since they only accounted for two per cent of the sample. The three main language groups where then compared in terms of their privacy concerns. The results of the MANOVA, univariate analysis and *post hoc* comparisons are shown in Table 7.36.

Table 7.36 Mean values and MANOVA results for different language groups

PRIVACY CONCERNS / LANGUAGE GROUPS	Privacy protection	Information misuse	Solicitation	Government protection	F value	p-value
English	4.8^{a}	3.9^{ab}	3.8^{a}	4.5		
Afrikaans	4.8^{b}	3.6^{a}	3.6^{b}	4.6		
Black African	4.6^{ab}	3.6^{b}	3.2^{ab}	4.5		
Univariate analyses	**0.0002**	**0.0014**	**0.0000**	0.0611		
Wilks' lambda					6.34	**0.0000**

[a] and/or [b]: The results of the Scheffé *post hoc* tests are indicated with [a] and/or [b]. All mean values containing the same letters (for example, [a]) indicate that the groups differ significantly from one another. All mean values containing different letters (for example, an [a] or [b]) indicate that these groups do not differ significantly from one another.

The Wilks' lambda value indicates a significant difference (p=0.0000) between the main language groups and their overall privacy concerns, providing support for H_{7b}. The subsequent univariate analyses revealed that these differences were significant for the first three factors, namely, privacy protection, information misuse and solicitation. The Scheffé *post hoc* tests showed that both the English- and Afrikaans-speaking groups differed significantly from the Black African language group on the privacy protection as well as the solicitation dimensions. Both the English-speaking and the Afrikaans-speaking groups were more concerned, as indicated by the higher mean values for these two groups. There were also significant differences between the Afrikaans-speaking and the Black African language group compared to the English-speaking group in terms of the information misuse dimension. Here the English-speaking group was more concerned than the other two groups, which had lower mean values. The high mean values of all three groups on Factor 4 (government protection) indicated that all the groups had strong concerns about government protection resulting in no significant differences between the groups.

7.5.2.10 H_{7c}: There is a significant difference between consumers in terms of their levels of education and their privacy concerns

Question 63 in the questionnaire (see Appendix 1) gave respondents five options to classify their highest level of education. To simplify the analysis and the hypothesis testing, these five groups were converted into three subgroups representing low, medium and high level of education. Respondents who indicated that their highest qualification was 'lower than Grade 8' or 'up to Grade 10' were classified as the low educational group. Respondents with a Grade 12 qualification were classified as the medium educational group. Respondents who indicated that they have a degree/diploma or a post-graduate degree/higher diploma were labelled as the high educational group. Most empirical studies reported strong positive relationships between educational levels and privacy concerns (refer to Chapter 5, Section 5.4.7.3). However, the studies indicated different relationships between more and less educated

levels and their privacy concerns. Hypothesis 7c was put to the test by means of MANOVA, follow-up analyses and *post hoc* tests, as reported in Table 7.37.

Table 7.37 Mean values and MANOVA results for different educational groups

PRIVACY CONCERNS / EDUCATIONAL GROUPS	Privacy protection	Information misuse	Solicitation	Government protection	F value	p-value
Low education level	4.6[a]	3.4[a]	3.3[ab]	4.5		
Medium education level	4.7[b]	3.6[a]	3.6[a]	4.5		
High education level	4.8[a]	4.0[a]	3.7[b]	4.6		
Univariate analyses	0.0600	**0.0000**	**0.0000**	0.8008		
Wilks' lambda					7.01	**0.0000**

[a] and/or [b]: The results of the Scheffé *post hoc* tests are indicated with [a] and/or [b]. All mean values containing the same letters (for example, [a]) indicate that the groups differ significantly from one another. All mean values containing different letters (for example, an [a] or [b]) indicate that these groups do not differ significantly from one another.

The specified significance level of 2.5 per cent resulted in a p-value of 0.0000, indicating support for H_{7c}. This means that there is a difference between groups with different levels of education and their privacy concerns. To determine where the differences between the three groups lie, univariate analyses were conducted on the different dependent variables (factors). This revealed that the differences between the different educational groups were related to information misuse (Factor 2) and solicitation (Factor 3). The Scheffé *post hoc* tests showed that all three educational groups differed from one another in terms of information misuse, with the highly educated group being the most concerned and the less educated group being the least concerned. Solicitation concerns (Factor 3) differed for the least educated group on the one hand and the medium and high level of education groups on the other. Here both the medium and high level of education groups were more concerned than the low level of education group. The high mean values for all three groups on Factor 1 (privacy protection) and Factor 4 (government protection) indicated that all the groups felt very

strongly about these two dimensions, resulting in no significant differences between the groups.

7.5.2.11 *H_{7d}: There is a significant difference between consumers in terms of their employment status and their privacy concerns*

Answers to Question 64 in the questionnaire (see Appendix 1) were divided into two meaningful groups. All the respondents who indicated that they were employed full time, part time or self-employed were labelled as the 'employed group'. Respondents who were students, homemakers, pensioners, not employed or physically unfit for work were classified as the 'not-employed group'. Only a few international studies have investigated the relationship between employment status and privacy concerns, and most of them have not found any significant relationships (refer to Chapter 5, Section 5.4.7.5). Table 7.38 shows the results of the MANOVA results for Hypothesis 7d, testing for significant differences between consumers' employment status and their privacy concerns.

Table 7.38 Mean values and MANOVA results for different employment status groups

PRIVACY CONCERNS EMPLOYED GROUPS	Privacy protection	Information misuse	Solicitation	Government protection	F value	p-value
Employed	4.7	3.8	3.6	4.5		
Not employed	4.7	3.6	3.5	4.6		
Univariate analyses	0.4255	**0.0044**	0.0514	0.2471		
Wilks' lambda					3.61	**0.0065**

The MANOVA result showed significant differences between the employed and the not employed group in terms of their overall concerns. **The null hypothesis was thus rejected, as there was support for H_{7d}.** The follow-up univariate analyses revealed that the significant overall result was due to a significant difference on only one privacy

dimension, namely information misuse (Factor 2). Here the employed group was more concerned than the unemployed group regarding the misuse of their personal information. No significant differences were found on any of the other three privacy dimensions (or factors).

7.5.2.12 *H_{7e}: There is a significant difference between consumers in terms of their income levels and their privacy concerns*

Nine percent of the respondents refused to answer Question 65, relating to their personal total monthly income. The remaining 91 per cent of the respondents were classified into one of three different income level groups. Respondents earning less than R2 000 per month were classified as the 'low-income group'. Respondents earning between R2 001 and R6 000 per month were classified as the 'middle-income group', and respondents earning more then R6 001 were labelled the 'high-income group'. International studies have reported conflicting results regarding the relationship between consumers' income levels and their privacy concerns, with some studies finding no relationship between income and privacy concerns (refer to Chapter 5, Section 5.4.7.5). Results of the MANOVA, univariate analysis and post hoc comparisons are set out in Table 7.39.

Table 7.39 Mean values and MANOVA results for different income groups

PRIVACY CONCERNS INCOME GROUPS	Privacy protection	Information misuse	Solicitation	Government protection	F value	p-value
Low income level	4.6^{ab}	3.5^{a}	3.3^{a}	4.5		
Middle income level	4.8^{a}	3.7^{b}	3.5^{b}	4.6		
High income level	4.8^{b}	4.1^{ab}	3.8^{ab}	4.5		
Univariate analyses	**0.0121**	**0.0000**	**0.0000**	0.2916		
Wilks' lambda					6.53	**0.0000**

[a] and/or [b]: The results of the Scheffé *post hoc* tests are indicated with [a] and/or [b]. All mean values containing the same letters (for example, [a]) indicate that the groups differ

significantly from one another. All mean values containing different letters (for example, an [a] or [b]) indicate that these groups do not differ significantly from one another.

The Wilks' lambda value indicates a significant difference (p=0.0000) between the three income groups in terms of their overall privacy concerns, providing support for H_{7e}. The follow-up univariate analyses showed that these differences were significant for the first three factors namely privacy protection, information misuse and solicitation. The Scheffé *post hoc* tests revealed that both the middle and high-income groups differed significantly from the low-income group in terms of the privacy protection dimension (or factor). Here both the middle and higher income groups have higher privacy concern levels. The *post hoc* tests also indicated that both the low and middle income groups differed significantly from the high income group on the information misuse as well as the solicitation dimensions (or factors), with the high income group being the more concerned group in both cases. **From results in Table 7.39 it can be concluded that higher income groups tend to have higher privacy concerns.**

7.5.2.13 *H_{7f}: There is a significant difference between males and females in terms of their privacy concerns*

Several international studies maintained that gender was strongly associated with privacy concerns (refer to Chapter 5, Section 5.4.7.6). It seems that generally, females express more concern about threats to their personal privacy than males do. The mean values of the two gender groups and the MANOVA results of Hypothesis 7f are shown in Table 7.40.

Table 7.40 Mean values and MANOVA results for different gender groups

PRIVACY CONCERNS GENDER GROUPS	Privacy protection	Information misuse	Solicitation	Government protection	F value	p-value
Male	4.7	3.7	3.5	4.4		
Female	4.8	3.7	3.7	4.6		
Univariate analyses	**0.0138**	0.5683	0.0481	**0.0005**		
Wilks' lambda					3.81	**0.0046**

The Wilks' lambda value indicates a significant difference (p=0.0046) between males and females in terms of their overall privacy concerns, providing support for H_{7f}. The subsequent univariate analyses revealed that these differences were significant for the first and the fourth privacy dimensions (or factors), namely privacy protection and government protection. In both cases females had higher privacy concerns, expressing their need for protection.

A summary of all of the above-discussed hypotheses is presented in Table 7.41, indicating whether support was found for the hypotheses.

Table 7.41 Summary of hypotheses tested

Alternative hypotheses		Supported or not supported
H_1	There is a significant difference between consumers in terms of their protective behaviour and their privacy concerns	Supported
H_{2a}	There is a significant difference between consumers who have been victims of invasions of privacy and consumers who have not been victims of invasions of privacy in terms of their privacy concerns	Supported
H_{2b}	There is a dependency between being a victim of invasion of privacy and gender	Supported

H_{3a}	There is a significant difference between consumers in terms of their level of awareness of name removal procedures and their privacy concerns	Supported
H_{3b}	There is a dependency between the level of awareness of name removal procedures and age	Not supported
H_{3c}	There is a dependency between the awareness of name removal procedures and levels of education	Not supported
H_4	There is a significant difference between Internet users and Internet non-users in terms of their privacy concerns	Supported
H_5	There is a significant difference between direct shoppers and non-direct shoppers in terms of their privacy concerns	Supported
H_6	The proportion of South African consumers is not equally represented in the different privacy segments.	Supported
H_{7a}	There is a significant difference between young and old people in terms of their privacy concerns	Supported
H_{7b}	There is a significant difference between the main language groups in terms of their privacy concerns	Supported
H_{7c}	There is a significant difference between consumers in terms of their levels of education and their privacy concerns	Supported
H_{7d}	There is a significant difference between consumers in terms of their employment status and their privacy concerns	Supported
H_{7e}	There is a significant difference between consumers in terms of their income levels and their privacy concerns	Supported
H_{7f}	There is a significant difference between males and females in terms of their privacy concerns	Supported

The above table provides a summary of the main findings in support of the secondary objectives specified in Chapter 5. Several other conclusions can be drawn from the data analyses conducted and discussed in this chapter. Some of the findings, in support of the primary objective, namely, to identify and explore the information privacy concerns of South African consumers, are summarised below:

- The exploratory factor analysis identified four underlying information privacy dimensions, namely privacy protection, information misuse, solicitation and government protection.
- Men seemed to be more likely than women to perceive themselves as victims of privacy invasion.
- Consumers who exercised full protective behaviour or limited protective behaviour were more concerned about privacy protection, information misuse and solicitation than consumers who did not exercise any protective behaviour.
- Consumers who had been victims of privacy invasions had higher information misuse and solicitation concerns than consumers who had not been victims of privacy invasions.
- Consumers who were not aware of any name removal options were more concerned about the misuse of their information, solicitation practices and government's protection than consumers who were aware of name removal options.
- Consumers' awareness of name removal options was not related to their levels of education or their age.
- Consumers who had undertaken Internet transactions were more concerned about the misuse of their information than consumers who had not used the Internet for transactions.
- Consumers who had not purchased directly in the past year showed higher concerns regarding the solicitation practices of companies than consumers who had purchased directly during the past year.
- Older consumers (40+ years) were more concerned than younger consumers (below 40 years) about the protection of their privacy, the misuse of their information and the solicitation practices of companies.
- English- and Afrikaans-speaking consumers were more concerned about privacy protection and solicitation practices than Black African language consumers.
- English-speaking consumers were more concerned about the misuse of their information than Afrikaans-speaking and Black African language consumers.

- Consumers with a high level of education (tertiary education) had higher privacy concerns regarding the misuse of their information compared to consumers with a medium or low level of education.
- Consumers with a medium or high level (Grade 12 and higher) of education were more concerned about companies' solicitation practices than consumers with a low level of education.
- Consumers who were employed showed higher concerns relating to information misuse than consumers who were not employed.
- Higher income consumers were more concerned regarding their privacy protection, information misuse and feelings toward solicitation than middle and low income consumers.
- Females were more concerned than males about the protection of their privacy by companies and government.
- Almost one-third of the respondents belonged to the Privacy Fundamentalist segment, which suggested that there are very high information privacy concerns among South Africans. The majority of South African consumers (58 per cent) belong to the Privacy Pragmatist segment, indicating balanced information privacy concerns, with only 11 per cent belonging to the Privacy Unconcerned segment, where there are very low levels of concern or no concern at all.

7.6 SUMMARY

In this chapter, the empirical results of the study were presented. First, there was a focus on the descriptive statistics, after which attention was given to the results of the exploratory factor analysis and the confirmatory factor analysis. The empirical analyses indicated that the information privacy scale used in this study was both reliable and valid. Finally, the empirical results were assessed against the formulated hypotheses, concluding with a summary of the outcomes of each hypothesis test. One conclusion that can be drawn from this chapter is that South African consumers show definite information privacy concern, and that their level of concern is very high in certain areas. In the final chapter, the above findings are interpreted, with particular reference to their

implications for marketers. Chapter 8 sets out conclusions and recommendations based on the main findings represented in this chapter.