

Chapter 4: Analysis of the data

We have seen in Chapter 3 how we went about organising our project and collecting the data for our analysis. In this chapter, we will be taking a closer look at the kind of data that we collected, and the statistical tests that we conducted in order to test our hypotheses about the determinants of a customer's intention to pay for an online content service, and to test whether the kind of paid content consumed in the past affects the user's intention.

4.1 Data Descriptive

The table below gives us a quick snapshot of the data collected through our survey:

Table 4.1: Grouping of respondents based on control questions

Variable	Category	Frequency	Percentage
Do you work in publishing?	Yes	52	24.2
	No	163	75.8
Do you work/study in the UK?	Yes	124	57.7
	No	91	42.3
Have you ever paid for content online?	Yes	173	80.0
	No	43	20.0

Source: SPSS Output

4.2 Data Reliability and Validity

The Cronbach Alpha co-efficient was used as an indicator of the reliability of the scales, and the results obtained were as follows:

Table 4.2: Values of Cronbach's α

Construct	Name of item	Number of items	Values of Cronbach's α
Perceived consequences	Con	4	.765
Perceived ease of use	Eou	2	.633
Social factors	Soc	2	.647
Satisfaction	Sat	3	.636
Alternatives	Alt	3	.642
Intention to use fee-based content services	Int	2	.691

Source: SPSS Output

The values of Cronbach's α are all above .6 and close to .7 for all of the constructs except for perceived consequences. We will accept these scales as being reliable given the proximity of Cronbach's α value to .7.

In this project, we are trying to find a correlation between the various dependent variables (perceived consequences, perceived ease of use, satisfaction, value placed compared to alternatives and social factors) and the independent variable (intention to pay). These are our inferential statistics (Sekaran, 2003 p.314). Carrying out a normality distribution check on our data, we see from the normal Q-Q plots (Figure (a) in Appendix 1) that all of the constructs are reasonably – if not very – normally distributed. We will thus consider our data to be normal and will be referring to parametric tests suitable for normally distributed data for our statistical analyses.

4.3 Preliminary Analyses

First, in order to establish whether or not there is a correlation between the intention to pay and the hypothesised determinants, a correlation test was carried out which gave us a measure of the relationship between the constructs and the dependant variable along with a direction i.e. whether the relationship is positive or negative. The table below gives us a value for the Pearson correlation coefficient – which gives us an indication of the strength of the relationship between the respective independent variable and the total intention. The Sig (2-tailed) value tells us whether the strength of this observed correlation is statistically significant.

Table 4.3: Pearson Correlation coefficients for constructs

		Total Intention
Total Perceived Consequences	Pearson Correlation	.656**
	Sig. (2-tailed)	.000
Total Perceived ease of use	Pearson Correlation	.197**
	Sig. (2-tailed)	.004
Total Social factors	Pearson Correlation	.436**
	Sig. (2-tailed)	.000
Total Alternatives	Pearson Correlation	.592**
	Sig. (2-tailed)	.000
Total Satisfaction	Pearson Correlation	.702**
	Sig. (2-tailed)	.000

** . Correlation is significant at the 0.01 level (2-tailed).

Source: SPSS Output

From the table above, we can see that there is a significant positive correlation of the intention with all of the determinant constructs.

The factors which come out as having the strongest link to intention are the satisfaction as a result of consuming a piece of content ($r=.702$), the perceived consequences

($r=.656$) and the perceived value of paid content compared to free alternatives ($r=.592$). There is a slightly weaker (though still good) correlation between social factors and intention ($r=.436$). As for perceived ease of use, the correlation is weak ($r=.197$), even though it is statistically significant. Table (b) in Appendix 1 shows the correlation coefficient between each of the constructs. However, this still does not tell us whether the independent variables have an causal effect on the dependent variable.

Next, we carried out an independent samples t-test (please see Appendix 2 for more detail) to see if the mean of intention was any different between respondents who had previously paid for online content, compared to those who had never paid for online content below. We see that there is a significant difference in the measure of total intention between those who have paid for online content before (Mean: 5.46, SD: 1.50), as compared to those who have not had the experience of paying for online content before (Mean: 6.34, SD: 1.57) – the value of Sig (2-tailed) is .001 which indicates a statistically significant difference in the total measure of intention between the two groups being studied. We will see in the multiple regression test results a further breakdown of this statistic to get a clearer idea of where exactly this difference lies.

An independent samples t-test was also carried out to gauge the difference in mean intention between respondents from the UK and the US – however, from the value of Sig (2-tailed), which is .833, we see that there is no statistically significant difference in the mean intention between the two groups.

4.4 Test of hypotheses

We already established that there are some significant correlations between our selected determinants and our dependent variable. In order to test our hypotheses, we need to establish that these determinants actually play a role in influencing the user's intention to pay for online content.

In order to do this, a multiple regression was carried out in order to see which determinants played a role influencing the user's intention, and also to what extent.

Tolerance figures for all constructs are above .10 and VIF values are well below 10, and we can be confident that multiple correlations with other constructs are very low and that multicollinearity is not a problem (Pallant, 2007 p.155).

In the table below, the value of Adjusted R Square tells us how much of the variance in the total intention is explained by our constructs. In percentage terms, this means that our model can account for 61.6% of the variance in total intention.

Table 4.4: Results of Multiple Regression Test

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.791 ^a	.626	.616	.98775

a. Predictors: (Constant), t_alt, t_eou, t_soc, t_con, t_sat

b. Dependent Variable: t_int

Source: SPSS Output

Looking at the ANOVA table below, we can get an understanding of the statistical significance of the above result.

Table 4.5: ANOVA Results for the Multiple Regression test

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	289.235	5	57.847	59.291	.000 ^a
	Residual	172.688	177	.976		
	Total	461.924	182			

a. Predictors: (Constant), t_alt, t_eou, t_soc, t_con, t_sat

b. Dependent Variable: t_int

Source: SPSS Output

The significance value being .000 indicates that this result is statistically significant. We thus conclude that our determinants – in combination at least – can influence a user’s intention to pay for content online.

In order to see how each individual determinant contributes to this influence, we look at the table of coefficients below.

Table 4.6: Coefficient values from Multiple Regression test

Model	Standardized Coefficients	Sig	Collinearity Statistics			
			Beta	Partial	Part	Tolerance
(Constant)		.008				
t_con	.327	.000	.391	.260	.633	1.579
t_eou	.070	.136	.112	.069	.958	1.044
t_soc	.137	.008	.198	.123	.816	1.225
t_sat	.361	.000	.377	.249	.475	2.106
t_alt	.144	.021	.172	.107	.550	1.819

Source: SPSS Output

The Beta value gives us an indication of the contribution of each determinant. The largest unique contribution comes from the satisfaction, followed very closely by the perceived consequences. Alternatives and social factors make much less of a contribution, with the perceived ease of use contributing very little. For all of the constructs except for the perceived ease of use, the significance levels are below 0.05, indicating that the construct is making a statistically significant unique contribution towards determining the user's intention.

If we square the "Part" values for each construct, we can get a percentage contribution to R Square uniquely for each construct, removing any shared contribution with the other constructs. We thus get perceived consequences alone contributing 6.8%, social factors contributing 1.5%, satisfaction contributing 6.2% and alternatives contributing 1.1% to the R Square value.

Going back to our hypotheses, we can conclude the following:

- *Hypothesis HA1: The greater the value of the perceived consequences of using the fee-based content, the more likely the customer's intention to adopt it.*

From Table 4.6 ($\beta = .327$, significance .000), we accept this hypothesis.

- *Hypothesis HA2: The greater the perceived ease of use of the fee-based online content, the more likely the customer's intention to adopt it.*

From Table 4.6 ($\beta = .070$, significance .136), we reject this hypothesis. In fact, the perceived ease of use does not have a direct, positive correlation with the customer's intention to pay for an online content service.

- *Hypothesis HA3: The higher a user perceives social influence in using the fee-based online content, the more likely their intention to use it.*

From Table 4.6 ($\beta = .137$, significance .008), we accept this hypothesis, keeping in mind the relatively small contribution of this determinant.

- *Hypothesis HA4: The greater the perceived value of using the fee-based online content as compared to available alternatives, the more likely the customer's intention to use it.*

From Table 4.6 ($\beta = .144$, significance .021), we accept this hypothesis.

- *Hypothesis HA5: The higher the level of satisfaction felt by the user after using the fee-based online content, the more likely their intention to adopt it.*

From Table 4.6 ($\beta = .361$, significance .000), we accept this hypothesis.

The last hypothesis which we are yet to investigate involves the study of the user's intention to pay for online content based on their experience of online content consumption.

We have categorised our sample into four groups:

- Those who have never paid for content online
- Those who have paid for content for personal/entertainment purposes as well as for professional/academic purposes
- Those who have only paid for content for entertainment purposes
- Those who have only paid for content for work/academic purposes

The separation into groups was possible because of the control questions which directly asked them if they had ever paid for a particular kind of content.

What we are trying to achieve here is an understanding of whether a user's past experience in online content consumption has an effect on their intention to pay for online content in the future.

For this, we ran an ANOVA test between all of our item scores against the four groups. We find that there is a statistical difference only between those who have never paid for any kind of content online, and those who have paid for online content for both entertainment and professional/academic purposes.

The hypothesis in question stands as below:

- *Hypothesis HB: A user's behaviour towards the online content service will differ, based on the type of online content.*

From the ANOVA test that we have done for groups, we find a pattern arising. Tables from the ANOVA test result can be found in Appendix 4, an interpretation of the results from the test are presented briefly below:

In the ANOVA table, the significance value is below .05 for four of the items, indicating a difference in measure between our four groups. Further, for all of these four items, the difference is only between respondents who have never paid for content before, and those who have paid for both entertainment/leisure as well as professional/information content.

We see that there is only a statistically significant difference between groups that have paid for both kinds of content in the past – and those who have never paid for content before. There does not seem to be any significant difference based on whether users have consumed entertainment/leisure or professional/information content.

Thus, we reject our hypothesis that the measure of intention differs based on the kind of content. However, it is an interesting outcome that the measures of perceived consequences, satisfaction and intention are affected by a user's past experience with online content.

The results that we have seen in this chapter open up this discussion about paid online content to a number of other issues in terms of implications and angles to explore further. Some of these issues are discussed in the following chapter.