Understanding channel purchase intentions

Broekhuizen, T.L.J.

6 Results and Discussion Study 1

This chapter presents the results of the first study according to the data analysis procedure outlined in the previous chapter. This study analyzes the consumers' motivations to shop online or offline at a specific multichannel bookseller (see section 5.3). First, the data collection is discussed, followed by a description of the characteristics of the online and offline buyers. Next, the stages identified in the former chapter are followed, ending in a discussion of the major findings.

6.1 Data collection

To ensure enough variation in channel purchase intentions, both consumers who are more likely to shop online and those who are more likely to shop offline were considered. As such, this study dealt with online and offline consumers of books. A well-known Dutch multichannel bookseller provided a sample of consumers who had recently purchased at least one book through their websites (online buyers). Next, a sample of offline consumers (offline buyers) was selected that only shopped through the offline stores of this multichannel bookseller.

Students collected the data for the offline sample. Three outlets of the multichannel bookseller were visited to increase the representativeness of the sample. Each outlet was visited twice (once during the week, once in the weekend) to reduce the effects of the days of the week. Based on interviews with managers, it became clear that weekend shoppers differed from those that shopped on weekdays in terms of socio-demographics. The people that visited the stores during the week tended to be older, and were more willing to spend additional time to shop than those in the weekend. To increase the likelihood that respondents had a real option to choose between channels, offline respondents were only

selected when they had access to the Internet, had not shopped through the website of the bookseller, but were familiar with the presence of the website. In three physical stores across the Netherlands 415 offline shoppers filled in the questionnaire of which 412 were usable. The data were collected during June 2004.

The data collection for the online sample (i.e. sample of customers who bought something via the website) was performed through an online survey. Online surveys were mainly chosen because of lower costs, faster response, and convenient collection (e.g. Ilieva, Baron and Healy 2002). Due to their experience with the Internet, online buyers were expected to fill in the online questionnaire without any trouble. However, one of the main concerns with online questionnaires is whether the data collected from online surveys are equivalent to (or comparable with) data collected from traditional mail surveys. Online surveys generally have a lower response rate than traditional mail surveys (Ilieva et al. 2002). To increase the response rate, a pocketbook was sent as an incentive to each respondent who filled in his/her address. Next, respondents may –probably due to their anonymity– respond with more extreme answers (Schaefer and Dillman 1998). Deutskens, De Ruyter and Wetzels (2004) tested the equivalence of online and offline surveys through a series of measurement invariance tests and concluded that both data methods were equivalent. This study also assesses the degree to which the instrument is invariant across contexts.

An email was sent to invite 1,019 shoppers who had bought a book through the website of a multichannel bookseller. The email address of the multichannel bookseller was used as sending address to evoke feelings of familiarity and reduce privacy concerns. The e-mail addressed the purpose of the questionnaire, the length of the questionnaire, the incentive, and guaranteed that the data would be treated confidentially. A link to a URL address was provided with instructions to start the questionnaire. In order to ensure that respondents answered the questionnaire only once, they had to log in with their email address and a generated password, which was provided in the email; it was not possible to log in twice with the same email address. A total of 241 (23.7%) questionnaires were filled in of which 239 were usable. The data were collected during June and July 2004. The response in both samples was deemed sufficient for subsequent analyses.

6.2 Respondent characteristics

Table 6.1 summarizes the respondent characteristics of the offline and online sample. As mentioned before, the online sample refers to the online buyers that shopped at least once through the website, whereas the offline sample refers to the offline buyers that did not shop through the website of this particular bookseller.

Table 6.1: Profile of the respondents for the online and offline sample

Socio-demographic variables		**Offline sample**		**Online sample**	
Gender	Male	218	52.9%	115	48.5%
	Female	194	47.1%	122	51.5%
Age	<19 years	34	8.4%	5	2.1%
	19-25 years	150	37.1%	51	21.5%
	26-40 years	119	29.5%	94	39.7%
	> 40 years	101	25.0%	87	36.7%
Income p.a.	Less than € 20,000	125	33.0%	42	18.9%
	€ 20,000 – €29,000	75	19.8%	60	27.0%
	€ 29,000 – €43,500	67	17.7%	54	24.3%
	€43,500 – €58,000	46	12.1%	31	14.0%
	€58,000 – € 72,500	28	7.7%	14	6.3%
	€72,500 or more	37	9.8%	21	9.5%
Education	Primary education	9	2.2%	2	0.8%
	Secondary education	25	7.1%	13	5.5%
	College	68	16.5%	34	14.4%
	Graduate	300	72.6%	171	72.1%
	Other	11	2.6%	17	7.2%

Both samples were composed of well-educated respondents with a relatively high income[13]. The majority of the respondents had a graduate background. The genders were equally distributed and were similar across contexts ($\chi^2(1)=1.16$ $p>.10$). The average age of the respondents for the online sample was slightly higher than that of the offline sample ($X_{online}=35.07$ years versus $X_{offline}=31.82$ years, $p<.05$). The income levels were not similar across the two samples ($\chi^2(5)=16.90$ $p<.01$), but no clear pattern could be found. Finally, the education levels appeared similar across contexts ($\chi^2(4)=9.21$ $p>.10$). The typical online shopper –male, high income, well educated, between 30 and 40 years old– does not clearly stand out[14].

The online questionnaire entailed fewer background-related questions. As such, the use of multiple channels could only be investigated for the offline sample. A total of 353 respondents indicated that their last purchase was made in the offline context. From these respondents, 15.4% indicated that they used the Internet prior to their *offline* purchase. This percentage is somewhat less than the 22% found in a large multichannel study for department, apparel and leisure stores (Bizrate.com 2001). The multichannel respondents used the Internet to search for specific book content (56%), price (42%), book availability (36%), background information (30%), and to get inspiration (16%). These respondents engaged in rather goal-oriented online search behavior, as inspiration was mentioned less frequently than specific search actions (book content, price). The Internet played a reasonable –but not substantial– role in their shopping process; an average of 3.93 was found on a scale from 1 (a marginal role) to 7 (a substantial role). In sum, the results suggest that consumers purposively decide to use the Internet for search activities prior to purchasing, and that this search activity is distinct from the decision to purchase through either channel. In this case, multichannel behavior is not expected to lead to blurred perceptions of shopping through either channel. The results also showed that online book purchases tended to be more goal-directed than offline book purchases. From the respondents who bought their last book online, 73.5% exactly knew what book to buy

[13] Based on a comparison with internal reports, it was found that the sample represented the bookseller's customer base. Generally, these high levels of education result in even higher income levels, but this moderate association can be explained as this bookseller has a relatively large student base.

[14] Although the typical online shopper did not clearly stand out, the results showed that younger men (age 26-40 years) were most positive towards the website.

prior to purchase, whereas only 49.4% of the people that bought their last book offline had a predetermined book in mind.

Some background information questions were asked to define the respondents' prior online shopping experience. In both samples, the majority of the respondents had experience in buying and/or ordering products or services online. From the offline sample, 61.5% indicated that they had shopped online for products or services *different* than books, whereas for the online sample, 72.4% had shopped online for products other than books (see Table 6.2). As expected, the online sample bought more frequently online than the offline sample based on a chi-square test ($\chi^2(5)=93.5$, $p<.001$). Yet, the majority of the respondents from the offline sample had at least shopped online once, indicating familiarity with the procedures involved with online shopping. Therefore, offline buyers were expected to be capable of conceiving the picture of shopping through the website of the multichannel bookseller.

Table 6.2: Prior online shopping experience (books excluded)

Total number of online purchases made	Offline sample N=408	Online sample N=239
0	38.5%	27.6%
1	17.6%	2.1%
2-3	20.3%	13.0%
4-6	12.3%	25.5%
7-10	7.1%	14.2%
> 10	4.2%	17.6%

6.3 Stage 1: Item analysis

Individual item analysis was performed to investigate the means and standard deviations of the items pertaining to the constructs in both the online and offline context for the base model. As respondents were asked to evaluate both channels, the results for online and offline buyers are shown separately here. Thus, the results are discussed for online buyers and offline buyers evaluating both the store and website. In stage 2 the scores of online buyers and offline buyers will be pooled for each context to investigate the

interrelationships between constructs (see section 5.4 for a discussion). Table 6.3 summarizes the main findings.

The initial pool of 28 pairs of scale items (see Table 5.1) was refined following generally accepted purification guidelines (e.g. Anderson and Gerbing 1982; 1988; Arnold and Reynolds 2003; Churchill 1979; Hair et al. 1998). As the goal was to have reliable and valid scales that apply to both contexts, the online and offline scales were examined simultaneously. When items performed poorly, they were removed simultaneously. First, corrected item-to-total subscale correlations were examined for each set of items representing a construct in the online and offline context. Items with corrected item-to-total subscale correlations below .50 were considered for deletion (Arnold and Reynolds 2003; Nunnally 1978). After investigation, four items (Risk2, Risk5, Time3, and Time4) were deleted from the online and offline sample. Second, correlations among items measuring the same dimension were examined. Items with inter-item correlations smaller than .40 with similar traits were considered for deletion. No additional items were removed in this phase. After these two item analyses, the remaining 24 pairs of items were used for further purification and refinement (see Table 6.3).

Table 6.3: Means, standard deviations, and mean differences

	Offline sample			**Online sample**		
Items[a,b]	Store[c,d]	Website[c,d]	Store-website[e]	Store[c,d]	Website[c,d]	Store-website[e]
SQ1	5.82 (1.02)	4.09 (1.13)	**1.73*** **	5.73 (1.14)	5.36 (1.33)	**.36*** **
SQ2	5.37 (1.12)	3.95 (.93)	**1.42*** **	5.31 (1.26)	4.93 (1.30)	**.37*** **
SQ3	6.08 (.97)	4.43 (1.13)	**1.65*** **	6.09 (.87)	5.85 (1.05)	**.24*** **
SQ4	5.97 (.97)	3.98 (1.06)	**1.99*** **	5.92 (1.14)	5.27 (1.38)	**.65*** **
SQ5	5.29 (1.23)	4.28 (1.04)	**1.01*** **	5.70 (1.24)	5.51 (1.32)	**.19***
Enjoy1	5.88 (1.26)	3.72 (1.41)	**2.15*** **	5.55 (1.50)	4.67 (1.58)	**.88*** **
Enjoy2	5.90 (1.14)	4.02 (1.26)	**1.88*** **	5.75 (1.22)	5.18 (1.42)	**.57*** **

Table 6.3: Means, standard deviations, and mean differences (continued)

Enjoy3	6.04 (1.15)	3.96 (1.37)	**2.09*** **	5.57 (1.21)	4.90 (1.42)	**.67*** **
Enjoy4	5.76 (1.27)	3.86 (1.40)	**1.89*** **	5.40 (1.51)	4.76 (1.67)	**.64*** **
Risk1	1.54 (1.20)	4.29 (1.79)	**-2.75*** **	1.71 (1.50)	3.16 (2.02)	**-1.46*** **
Risk3	1.88 (1.15)	3.87 (1.62)	**-2.00*** **	1.89 (1.01)	3.24 (1.27)	**-1.35*** **
Risk4	1.61 (.88)	3.93 (1.57)	**-2.31*** **	1.72 (1.12)	3.02 (1.65)	**-1.31*** **
Time1	3.73 (1.62)	3.05 (1.44)	**.67*** **	4.07 (1.70)	2.09 (1.19)	**1.71*** **
Time2	3.25 (1.56)	3.27 (1.46)	-.02	3.85 (1.74)	2.14 (1.20)	**1.98*** **
MQ1	5.95 (1.05)	4.85 (1.30)	**1.10*** **	5.92 (1.16)	5.99 (1.06)	-.07
MQ2	5.91 (1.08)	4.85 (1.32)	**1.06*** **	5.89 (1.15)	5.95 (1.09)	-.06
Price1[f]	3.75 (1.35)	3.77 (1.18)	-.03	3.18 (1.26)	3.31 (1.30)	-.13
Price2[f]	3.25 (1.20)	3.62 (1.09)	**-.37*** **	3.11 (1.29)	3.41 (1.32)	**-.30*** **
PV1	5.50 (1.25)	4.82 (1.20)	**.67*** **	5.79 (1.03)	5.71 (1.06)	.08
PV2	4.88 (1.31)	4.62 (1.13)	**.27*** **	5.24 (1.24)	5.21 (1.28)	.03
PV3	4.58 (1.35)	4.44 (1.17)	**.14***	4.90 (1.28)	4.87 (1.33)	.03
Int1	5.73 (1.33)	3.21 (1.70)	**2.52*** **	5.04 (1.65)	5.69 (1.34)	**-.64*** **
Int2	5.51 (1.32)	3.67 (1.46)	**1.84*** **	5.22 (1.49)	5.43 (1.51)	-.21
Int3	5.15 (1.50)	3.03 (1.50)	**2.11*** **	4.54 (1.73)	5.04 (1.66)	**-.49*** **

* p < .05; ** p < .01; *** p < .001

Notes:

a. SQ=Service quality; Enjoy=Enjoyment; Risk=Perceived risk; Time=Time/effort costs; MQ=Merchandise quality; Price=Monetary price; PV=Perceived value; Int=Purchase intentions.

b. Each item (e.g. SQ1) is measured in the offline and online context; a total of 24 pairs of items are represented. Reverse-scaled items (Price1, Price2, Time1 and Time2, see section 5.3) were recoded during data entry for consistency.
c. Item means are based on 7-point Likert scale (1=totally disagree, 7=totally agree).
d. Standard deviations are displayed between brackets.
e. Figures in bold represent significant mean differences measured through paired-sample *t*-tests. Sample sizes for the paired *t*-tests ranged from 399 to 405 for offline buyers, and from 209 to 231 for online buyers, because of missing data.
f. Price level refers to the end price consumers have to pay. Respondents were instructed to take into account the delivery costs.

All items in Table 6.3 indicate positive outcomes, except for items of price, time/effort, and risk. High scores on the latter items indicate that consumers endure higher prices, higher time and effort expenditures and more risk.

To investigate the tradeoffs consumers make, the differences between the online and offline perceptions were analyzed for offline and online buyers. Offline buyers generally find offline shopping to outperform online shopping. On two items (Time2, Price1) the mean differences are not significant, whereas on one item they indicate that online shopping requires less time and effort. Thus, although offline shopping requires somewhat more time and effort, offline buyers strongly prefer to shop offline. Online buyers are more positive towards shopping through the website of the bookseller; they, for example, find the online channel to strongly outperform the offline channel in terms of time/effort required. They tradeoff these time/effort savings against lower service levels, less enjoyment and more risk. The following results stand out. The perceptual differences between shopping online and offline are much smaller for online buyers relative to offline buyers. Although both groups agree that offline shopping offers better service, more enjoyment and less risk, the perceptual differences between online and offline shopping are much smaller for online buyers. This can be partly explained through a gain in experience. For instance, there was evidence that online risk perceptions reduce with increasing online experience (cf. Montoya-Weiss et al. 2003); the average of the three risk items decreased significantly ($F(5,617)=15.56$, $p<.001$) from 4.20 (zero purchases), 4.07 (1 purchase), 3.63 (2-3 purchases), 3.21 (4-6 purchase), 3.09 (6-10 purchases) to 2.89 (more than 10 purchases).

Offline buyers also indicate to receive more value for money[15] in the store, whereas online buyers are indifferent towards the value received online and offline. Finally, online buyers incline to shop through the website, whereas offline shoppers intend to remain loyal to the store. Although these scores provide useful insights into the relative performance of channels and the tradeoffs the groups of buyers make, they cannot be used to infer the relative importance of these factors in each channel.

6.4 Stage 2: Exploratory factor analysis

After the item analyses, the items were subjected to exploratory factor analysis with principal axis factoring and oblique rotation, with the scree test criterion to identify the number of factors to extract (Arnold and Reynolds 2003; Hair et al. 1998; Nunnally 1978). Oblique rotation was performed rather than using Varimax rotation. Varimax rotation would imply uncorrelated factors (Rossiter 2002), which was unlikely to be the case. In an iterative manner, a series of factor analyses was performed to eliminate items with low loadings (<.50), low communalities (<.30), and/or high cross-loadings (>.40) (cf. Churchill 1979; Hair et al. 1998; Rossiter 2002).

For the exploratory factor analysis, the datasets were pooled to infer the underlying structure of factors for each context (see section 5.4). In other words, online buyers and offline buyers were pooled with respect to their evaluations of a given channel. Due to the model complexity, it was decided that for each context two separate exploratory analyses were performed regarding the base model, namely: (1) antecedents of perceived value and intentions, and (2) perceived value and intentions.

6.4.1 Antecedents of perceived value and purchase intentions

For the exploratory factor analysis regarding the antecedents of perceived value, the pooled sample regarding the offline context consisted of 573 respondents (395 offline buyers, 178 online buyers). The online context consisted of 537 respondents (384 offline buyers, 153 online buyers).

[15] In Chapter 6 and 7 the terms "perceived value" and "value for money" are used interchangeably to indicate the *construct* of perceived value.

In the online context, price and assortment items loaded on the same factor. Based on the scree test criterion, a six-factor solution was found in both contexts. The factor solutions accounted for approximately 73.6% and 67.4% of the total variance for the offline and online context, respectively. The KMO measures of sampling adequacy were .852 for the offline context and .894 for the online context, supporting the use of factor analysis[16]. Table 6.3 shows the results for the offline and online context, respectively.

Table 6.3: Exploratory factor analysis for antecedents of perceived value and intentions

Offline context/ Online context N=573/N=537	1	2	3	4	5	6
SQ1	**.73/.75**					
SQ2	**.74/.86**					
SQ3	**.77/.70**					
SQ4	**.83/.78**					
SQ5	**.52/.64**					
Enjoy1		**.78/.80**				
Enjoy2		**.90/.78**				
Enjoy3		**.86/.94**				
Enjoy4		**.75/.78**				
Risk1			**.47/.62**			
Risk3			**.85/.86**			
Risk4			**.81/.84**			
Time1				**.85/.84**		
Time2				**.76/.94**		
MQ1					**.94/.93**	
MQ2					**.78/.60**	
Price1						**.79/.83**
Price2						**.54/.53**

[16] Hair et al. (1998) indicate that a KMO of .80 or above is meritorious; between .80 and .70 is middling; between .70 and .60 is mediocre; between .60 and .50 is miserable; below .50 is unacceptable.

Table 6.3: Exploratory factor analysis for antecedents of perceived value and intentions (continued)

Offline context/ Online context N=573/N=537	1	2	3	4	5	6
Cronbach's alpha	.86/.89	.90/.90	.71/.83	.79/.88	.88/.84	.65/.67
Eigenvalues	6.16/7.38	1.78/1.81	1.64/1.52	1.41/1.37	1.24/1.02	1.01/.86
Variance extracted	73.6% / 67.4%					
KMO measure	.852 / .894					

Pattern Matrix shown, Principal Axis Factoring, Oblique Rotation.
Note: The first figure refers to the offline context, the second figure to the online context. Loadings <.30 are not shown.

6.4.2 Perceived value and purchase intentions

For the exploratory factor analysis regarding the perceived value and purchase intentions, the pooled sample regarding the offline context consisted of 618 respondents (405 offline buyers, 213 online buyers). The online context consisted of 611 respondents (393 offline buyers, 218 online buyers).

In both the online and offline context a two-factor solution was found regarding offline and online perceived value and intentions, whereas three constructs (i.e. perceived value, perceived value competing channel, and purchase intentions) were anticipated. In both contexts, the items of offline and online perceived value loaded on the same factor. This unexpected finding can be explained as both online and offline perceived value are derived from the same bookseller; therefore, they are highly correlated. Moreover, consumers consider the value derived from competing alternatives, as perceived value is judged relativistically (Holbrook 1996). Based on the results, it was decided to perform subsequent confirmatory factor analyses without the value from the competing channel. The two-factor solutions clearly distinguished value perceptions from channel purchase intentions; the solutions accounted for approximately 76.2% and 80.4% of total variance for the offline and online context, respectively. The KMO measures of sampling adequacy supported the use of factor analysis (see Table 6.4).

Table 6.4: Exploratory factor analysis for perceived value and intentions

Offline context/ Online context N=618/N=611	**1**	**2**
PV1	**.63/.52**	
PV2	**.95/.96**	
PV3	**.85/.85**	
Int1		**.84/.91**
Int2		**.72/.86**
Int3		**.83/.87**
Cronbach's alpha	.85/.83	.83/.91
Eigenvalues	2.99/1.15	1.58/3.16
Variance extracted	76.2% / 80.4%	
KMO measure	.734 / .785	

Pattern Matrix shown, Principal Axis Factoring, Oblique Rotation.
Note: The first figure refers to the offline context, the second figure to the online context. Loadings <.30 are not shown.

After performing the exploratory factor analyses, 24 items were retained, which measured 8 constructs in each context. Cronbach alpha coefficients were used to assess scale reliabilities. The reliability coefficients ranged from .65 to .91 (see Table 6.3 and 6.4). Except for the two price constructs, all scales met the suggested minimum level for internal consistency of .70 (Nunnally 1978). These values suggested that the scales were reliable and could be used for further analysis.

6.5 Stage 3: Confirmatory factor analysis

After performing exploratory factor analyses, confirmatory factor analyses (CFAs) were conducted with AMOS 5.0 (Arbuckle and Wothke 1999). Following Anderson and Gerbing (1988), the measurement model (relationships between observed items and latent constructs) was analyzed before the structural model (relationships between latent constructs). The logic of this argument is that it is essential to understand *what* one is measuring prior to testing relationships (Vandenberg and Lance 2000). The CFAs were run including both the exogenous (antecedents of perceived value) and endogenous (perceived value and intentions) part without any structural relationships.

6.5.1 Model-fitting procedure

Before analyzing the measurement model, it is necessary to determine how to treat missing data. The most commonly used method for dealing with missing data is listwise deletion (Hair et al. 1998). This study also adopts this method for the offline context. For the online context, however, it was chosen to minimize the number of omitted online buyers by replacing missing values using an expectation-maximization (EM) procedure. Research has shown that the EM method introduces the least bias into the estimated models (Hair et al. 1998). After performing listwise deletion, 143 online buyers would be retained but by replacing the missing values of respondents who had only 1 missing value, 197 online buyers were retained[17]. Subsequently, the effect of replacing the missing values on the measurement model (i.e. online buyers regarding online context) was assessed by investigating the standardized loadings in the dataset with and without replaced missing values (Arbuckle and Wothke 1999). The standardized loadings appeared very similar, and it was decided that the influence of replacing missing data for online buyers regarding the online context was negligible.

As the proposed estimation technique, maximum likelihood (ML), assumes multivariate normality, skewness and kurtosis were investigated. The distributions of the pooled datasets (N=564 and N=579) showed no strong skewness and kurtosis in both datasets, and no adaptations were required.

The fit of the CFA models were assessed on a number of fit indices, including chi-square, relative chi-square, goodness-of-fit (GFI), adjusted goodness of fit index (AGFI), non-normed fit index (NNFI) (Hu and Bentler 1995), relative fit index (RFI) (Bollen 1986), comparative fit index (CFI) (Bentler 1990), standardized root mean square residual (SRMR), and root mean square error of approximation (RMSEA) (Bollen 1989). For a detailed discussion of these fit indices, see Bollen (1989) and Hair et al. (1998).

[17] The decision to impute missing values was made to enable CFAs for each group within each context with substantial sample sizes. This facilitates the test of the moderating effect of the degree of prior online experience in the online context, as the subsample sizes are close to or exceed the commonly required sample size of 200 (Hair et al. 1998).

The initial ML estimation test of the 24 items produced fit indices slightly below acceptable thresholds for both contexts (χ^2/df= 2.78/3.21, GFI=.89/.87, RMSEA=.058/.077). The model was consequently refined by eliminating items contributing most to lack of fit, as indicated by the standardized residuals and modification indices (Steenkamp and Van Trijp 1991). After this iterative process, four items (Risk1, SQ1, SQ3, and Enjoy4) were excluded to improve fit. The final model showed reasonable fit indices (see Table 6.5).

Table 6.5: Fit indices for the online and offline context

	Recommended level	Offline context N=564	Online context N=579
χ^2	-	328.56	390.77
df	-	142	142
χ^2/df	<3.00	2.31	2.75
GFI	>.90	.94	.94
AGFI	>.90	.92	.91
NNFI	>.90 (>.95)	.95	.95
CFI	>.90 (>.95)	.97	.96
RFI	>.90 (>.95)	.92	.93
SRMR	<.05	.045	.055
RMSEA	<.08 (<.05)	.048	.055

Note: GFI=Goodness of Fit Index; AGFI=Adjusted Goodness of Fit Index; PGFI=Parsimony Goodness of Fit Index; NNFI=Non-Normed Fit Index; CFI=Comparative Fit Index; RFI=Relative Fit Index; SRMR=Standardized Root Mean Residual; RMSEA=Root Mean Square Error of Approximation.

The chi-square statistic showed that the models were significant (p<.0001), indicating that the specification of the factor loadings, factor variances/covariances, and error variances for the models under study are not valid. However, this is not uncommon, as the chi-square statistic is sensitive to departures from multivariate normality and large sample sizes (Diamantopoulos and Siguaw 2000; Hair et al. 1998). Due to the sensitivity of the chi-square statistic, other overall measures have been proposed, such as the normed chi square (Byrne 2001); the ratios of the chi-square to the degrees of freedom were beneath the recommended level of 3.00 (Carmines and McIver 1981; Byrne 2001). The GFI and AGFI

exceeded the recommended level of .90 in both contexts (Byrne 2001; Hair et al. 1998). The CFI measures the relative improvement of fit of the hypothesized models compared with the independence model. Although a value of >.90 was initially considered representative of a well-fitting model (Bentler 1992), more recently a revised cutoff value close to .95 is recommended (Hu and Bentler 1999). Both models met this revised cutoff. For the SRMR, values below .05 are indicative of good fit, indicating that the online context just falls outside the recommended level (Byrne 2001; Diamantopoulos and Siguaw 2000). The RMSEA values of both models also showed acceptable fit (Diamantopoulos and Siguaw 2000; Hair et al. 1998). Based on this, it can be concluded that the hypothesized models fit the data well.

6.5.2 Assessment of convergent validity, discriminant validity and reliability

CFA tests were run to test the convergent and discriminant validity of the constructs in the base model; these two types of validity are often used to assess construct validity. Convergent validity assesses the degree to which two measures of the same construct are correlated (Hair et al. 1998). Convergent validity can be assessed from the measurement model by determining whether each item's estimated maximum likelihood loading on its assigned construct factor is significant (Anderson and Gerbing 1988). The assessment of the measurement properties of the scales indicated that the factor loadings were high and significant ($p < 0.001$) in both contexts, which satisfies the criteria for convergent validity (Hair et al. 1998). Additionally, the average variance extracted (AVE) exceeded the recommended level of .50 for all constructs except price, showing that the variance captured by constructs was larger than variance due to error (Fornell and Larcker 1981). Although the second price item (Price2) scored poorly in both contexts, it was decided to maintain the item to ensure content validity (Peter and Churchill 1986), and to account for measurement error by having at least two items per construct.

Although the constructs in this study are conceptually related, they should also possess discriminant validity. Discriminant validity assesses the extent to which a measure does not correlate with other constructs from which it is supposed to differ (Malhotra 1996). Discriminant validity was first assessed by checking whether the confidence interval (± two standard errors) for each pairwise correlation estimate did not include the value of 1.0

(Anderson and Gerbing 1988). Each of constructs satisfied this first criterion in both contexts (see Table 6.7). The second test of discriminant validity was performed by assessing whether fit was improved when any pair of constructs was collapsed into a single factor (De Haes et al. 2004). The results of χ^2 difference tests indicated that discriminant validity was upheld in all pairwise tests. The third test involved whether the squared correlation between two constructs exceeded the AVE for each of the two constructs. Only the two price constructs did not meet this criterion, when it was correlated with value for money.

The construct reliabilities demonstrated that the scales were reliable, as they met the minimum construct reliability of .60 (Bagozzi and Yi 1988). In sum, the measurement model showed evidence for convergent and discriminant validity, as well as reliability. As such, the structural analyses could be performed with some confidence. Table 6.6 shows the standardized loadings, average variance extracted, and construct reliabilities for each construct.

Table 6.6: Item loadings, construct reliabilities and average variance extracted

	Offline context N=564		**Online context N=579**	
	Standardized loading[a,b,c]	**AVE[d]**	**Standardized loading[a,b,c]**	**AVE[d]**
Service quality	**.76**	*.52*	**.81**	*.59*
SQ2: high-quality services[e]	.76		.78	
SQ4: willingness to respond	.73 (15.20)		.76 (17.37)	
SQ5: reliability/fulfillment	.67 (14.16)		.76 (17.32)	
Merchandise quality	**.89**	*.79*	**.89**	*.81*
MQ1: good selection	.89		.90	
MQ2: wide selection of interesting books	.90 (20.01)		.90 (23.12)	

Table 6.6 Item loadings, construct reliabilities and average variance extracted (continued)

Monetary price	**.66**	*.50*	**.66**	*.49*
Price1: low price level (r)	.81		.77	
Price2: attractive offers (r)	.58 (11.14)		.63 (11.60)	
Perceived risk	**.73**	*.58*	**.84**	*.73*
Risk3: purchasing uncertainty	.66		.76	
Risk4: things can easily go wrong	.85 (7.07)		.94 (13.15)	
Time/effort costs	**.79**	*.65*	**.84**	*.73*
Time1: shopping efficiency (r)	.77		.79	
Time2: requires not lot of time/effort (r)	.83 (8.22)		.91 (17.00)	
Enjoyment	**.89**	*.73*	**.89**	*.74*
Enjoy1: shopping is fun	.84		.82	
Enjoy2: shopping is enjoyable	.87 (24.21)		.86 (23.76)	
Enjoy3: shopping is interesting	.86 (23.79)		.90 (24.83)	
Value for money	**.86**	*.67*	**.84**	*.65*
PV1: value for money	.67		.66	
PV2: price/quality ratio	.90 (17.78)		.88 (17.03)	
PV3: get versus give	.88 (17.61)		.85 (16.80)	
Purchase intentions	**.84**	*.64*	**.91**	*.78*
Int1: shopping likelihood	.83		.90	
Int2: willing to recommend	.76 (18.41)		.87 (29.07)	
Int3: future purchase intent	.82 (19.73)		.87 (29.37)	

Notes:

a. Figures in bold represent construct reliabilities, which were calculated based on the formula provided by Hair et al. (1998, p. 624).

b. Figures between brackets represent *t*-values of the factor loadings. The first item of each construct was used as a reference item.
c. Based on one-tailed tests, *t*-values greater than 1.65 are significant at $p<.05$; *t*-values greater than 2.33 are significant at $p<.01$ (cf. Parasuraman et al. 2005).
d. The average variance extracted (AVE) was calculated based on the formula provided by Fornell and Larcker (1981).
e. For the exact wording of the items, see Table 5.1.

6.5.3 Assessment of correlations and multicollinearity

Table 6.7 shows the correlations among the latent variables to indicate the interrelationships for the online and offline context. Overall, the correlations are somewhat stronger in the online context, which seem to be the result of a greater dispersion of the scores; most respondents have favorable perceptions towards offline shopping, whereas the perceptions towards online shopping are more scattered. Apart from this, most correlations are moderately high (i.e. .40-.60 range). A number of correlations between the constructs are high (above .60). The highest correlations are between price and value for money (i.e. $\varrho_{offline}$= -.75 and ϱ_{online}= -.71). This is not surprisingly, as the value for money consumers receive from the bookseller is naturally strongly correlated with its price level[18]. The investigation of the correlations between the shopping experience costs and benefits showed that enjoyment was negatively correlated with perceived risk ($\varrho_{offline}$= -.28 and ϱ_{online}= -.45), time/effort costs were negatively correlated with enjoyment ($\varrho_{offline}$= -.26 and ϱ_{online}= -.45), and perceived risk and time/effort costs were moderately positively correlated in the online context (ϱ_{online}= .31), but not significantly correlated in the offline context ($\varrho_{offline}$= .09). The moderate correlations between the three types of shopping experience costs and benefits support the distinctiveness of these factors.

To check for multicollinearity, regression –with unweighted summated scores– and correlation analyses were performed (see Table 6.7). No multicollinearity problems were encountered since the largest variance inflation (VIF) value was 2.10, which was lower than the commonly suggested cut-off value of 10 (Hair et al. 1998), and the more restricted level of 2.5 (Allison 1999). Next, the highest pairwise correlation among the *independent* factors was highest for service quality and perceived value in the offline context (ϱ=.60).

[18] Baker et al. (2002) found a correlation of .64 between price and value perceptions.

According to Hair et al. (1998), correlations between independent variables of .90 and above indicate multicollinearity problems.

Table 6.7: Correlations between latent factors after CFA

	SQ[a,b,c]	MQ	Price	Risk	Time	Enjoy	PV	Int
Service quality		.54 (.05)	-.58 (.06)	-.45 (.05)	-.45 (.05)	.57 (.05)	.55 (.05)	.62 (.04)
Merchandise quality	.54 (.06)		-.48 (.05)	-.38 (.05)	-.56 (.04)	.47 (.04)	.46 (.05)	.55 (.04)
Monetary price	-.51 (.07)	-.45 (.06)		.28 (.05)	.38 (.06)	-.38 (.06)	-.71 (.05)	-.42 (.05)
Perceived risk	-.34 (.07)	-.28 (.06)	.30 (.06)		.31 (.05)	-.42 (.05)	-.32 (.05)	-.45 (.04)
Time/effort costs	-.30 (.06)	-.23 (.06)	.15 (.07)	.09 (.06)		-.45 (.05)	-.47 (.05)	-.54 (.04)
Enjoyment	.56 (.05)	.50 (.05)	-.25 (.07)	-.28 (.06)	-.26 (.06)		.38 (.05)	.66 (.04)
Perceived value	.60 (.04)	.41 (.06)	-.75 (.05)	-.21 (.05)	-.18 (.06)	.33 (.06)		.45 (.05)
Purchase intentions	.47 (.06)	.43 (.05)	-.33 (.07)	-.29 (.06)	-.35 (.06)	.53 (.06)	.34 (.07)	

Notes:

a. SQ=Service quality; MQ=Merchandise quality; Price=Monetary price; Risk=Perceived risk; Time=Time/effort costs; Enjoy=Enjoyment; PV=Perceived value; Int=Purchase intentions.

b. Correlations offline context below diagonal, correlations online context above diagonal.

c. Standard errors are displayed between brackets and were derived by bootstrapping with 500 replications.

6.6 Stage 4: Multiple group CFA for base model

As a prerequisite to testing for differences in the strength of structural relationships, it is customary to first establish a baseline model for each context separately (Byrne 2001, p. 175). The baseline models are also used to test the first hypotheses based on the significance of the structural relationships.

6.6.1 Assessment of structural relationships for base model

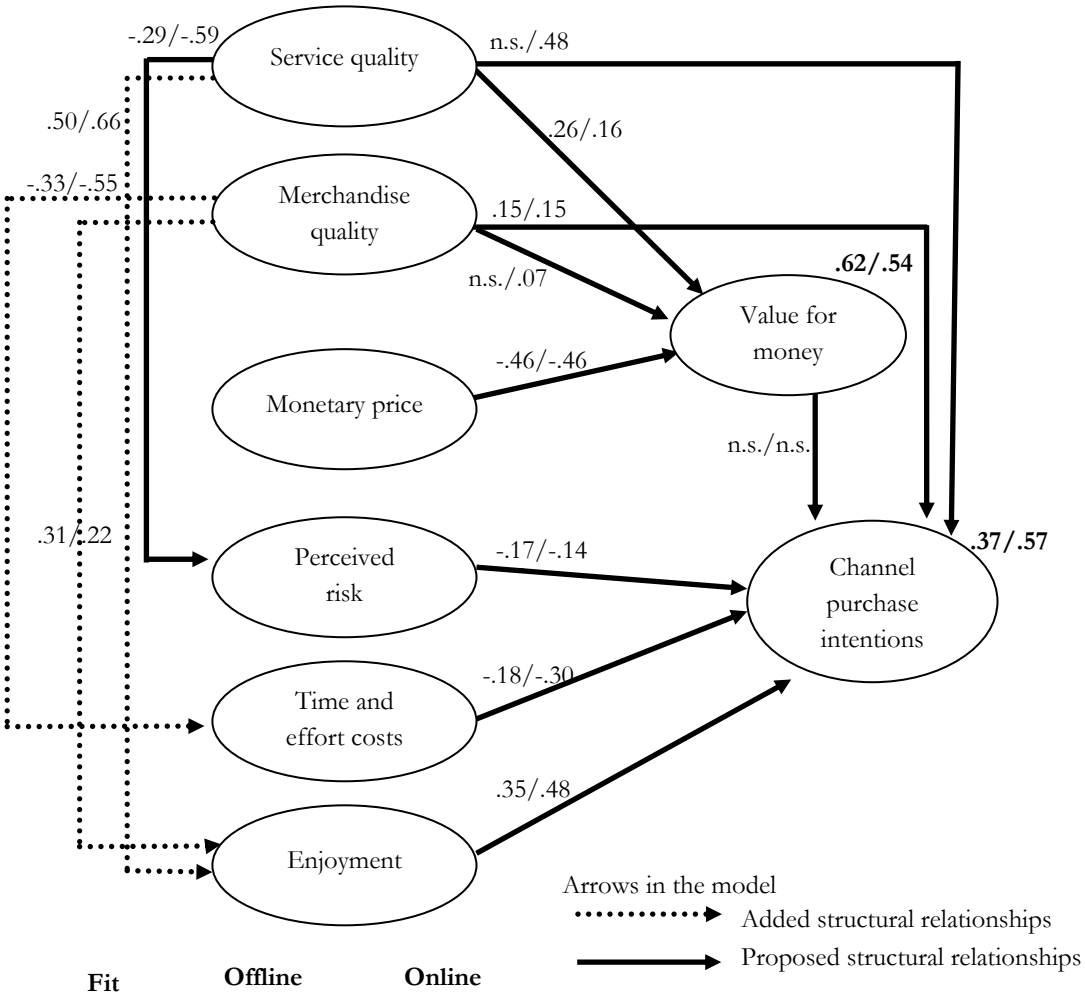
The proposed structural model (see conceptual model in Chapter 4) showed less than acceptable fit indices for the offline/online context (χ^2/df= 3.67/5.41, GFI=.91/.90, SRMR=.15/.20, RMSEA=.069/.087). Model respecifications were considered to improve fit, based on the examination of the normalized residuals and the modification indices (Hair et al. 1998). The purpose of these respecifications is to achieve the most parsimonious model that provides the best fit among the alternatives (Byrne 2001). In this study, modifications were only performed by adding new relationships to the model that were suggested by modification indices[19]. None of the insignificant relationships were removed, as they might slightly alter the other relationships, making a side-by-side evaluation less amenable. Particularly, in the case that one relationship turns out to be significant in one context and insignificant in the other, the researcher needs to allow for different path diagrams for different contexts. Although AMOS 5.0 enables this option (Byrne 2001), the strength of each relationship can be best compared when the models are identical[20]. The influence of maintaining insignificant paths was, however, marginal, as subsequent analyses showed no virtual differences in the strength of the other (significant) relationships, when the insignificant relationships were maintained.

The number of model modifications should be kept low, and only those that correct for relatively severe problems of model fit should be introduced (MacCullum et al. 1992; Steenkamp and Baumgartner 1998). The analyses showed that three additional relationships were required to reach acceptable fit indices. These three relationships were added in each context based on statistical and theoretical arguments (Byrne 2001). Merchandise quality had a significant relationship with both enjoyment and time/effort costs. When perceptions of merchandise quality increase, it is likely that consumers find the process more enjoyable as they can browse through store/website to find their desired books. Consumers, particularly those who approach retail environments to browse (Bloch, Sherrel and Ridgway 1986), often have a desire for the experiential aspects of shopping (Mathwick

[19] To avoid triviality, no correlations between error terms were allowed. As noted by Jöreskog (1993, p. 297) "every correlation between error terms must be justified and interpreted substantively." The justification for these correlations is often arbitrary.
[20] This suggestion was provided by Dr. Boomsma (University of Groningen, Department of Sociology).

et al. 2002). For them, better assortments will yield more enjoyable shopping experiences. Simultaneously, improvements in merchandise quality lead to time and effort savings, as consumers can find their book(s) of interest more quickly (cf. Szymanski and Hise 2000). Assortments of low quality, which can be the result of unstructured displays and routings or out-of-stock situations, will require more time and effort from consumers to find their desired book. Finally, an additional relationship was needed between service quality and enjoyment. Apart from the functional benefits, service quality also leads to emotional/affective benefits. In this respect, empathy, courtesy, problem solving and showing sincere interest in fulfilling the individual consumer needs are elements of (retail) service quality (cf. Dabholkar et al. 1996; Parasuraman et al. 1985; Wolfinbarger and Gilly 2003) that generally result in more enjoyment. Additional services are often aimed at facilitating the shopping process and taking away frustration (i.e. reducing the cognitive efforts), which makes it more amenable that consumers find the shopping process enjoyable. The robustness of the additional relationships will be tested in the replication study (see Chapter 7).

The final model is graphically displayed in Figure 6.1.



Fit indices	Offline context	Online context
χ^2	365.10	466.58
Df	154	154
χ^2/df	2.37	3.03
GFI	.94	.92
AGFI	.92	.91
NNFI	.95	.94
CFI	.96	.95
RFI	.92	.92
SRMR	.055	.074
RMSEA	.049	.059

Figure 6.1: Coefficients for the offline/online context for base model

Notes: The unstandardized structural coefficients are displayed for the offline/online context. Figures in bold represent the percentage of explained variance in the endogenous variables. N.s. represents coefficients that are not significant from zero at a .05 significance level based on one-tailed tests.

After the modifications, the models showed good fit indices for the offline/online context (see above). Moreover, the results generally support the hypothesized model. A considerable part of the variance of the endogenous factors were explained by their predictors –based on the squared multiple correlations (SMCs) (Byrne 2001; Hair et al 1998). For the offline context, the predictors account for 37.4% of the variance associated with purchase intentions, whereas in the online context 56.5% is explained. Next, the predictors to a large extent explained variations in offline perceived value (62.1%) and online perceived value (53.9%). Table 6.8 and 6.9 show the unstandardized and standardized structural relationships and their *t*-values. The unstandardized structural coefficients (also known as unstandardized structural weights) are similar to the regression weights in regression and are comparable across samples, as they are measured in their original metric (Diamantopoulos and Siguaw 2000; Hair et al. 1998). Standardized structural coefficients lose their natural meaning as they are cast in correlation rather than in covariance terms, but they help identifying the relative contribution of independent latent factors in influencing the endogenous latent factors (Diamantopoulos and Siguaw 2000). Standardized structural coefficients cannot easily be compared across samples. Bagozzi (1980, p. 187) warns, "standardized parameters are appropriate only when one desires to compare the relative contribution of a number of independent variables on the same dependent variable and for the same sample of observations. They are not appropriate and can lead to erroneous inferences when one wishes to make comparisons across populations of samples." For example, an inspection of the standardized structural coefficients of the determinants of offline purchase intentions (see Table 6.8) shows that enjoyment (.32) has the strongest direct effect on purchase intentions, followed by time/effort costs (-.20), merchandise quality (.12), and perceived risk (-.10). The effect of enjoyment on offline purchase intentions is three times as large as the effect of perceived risk. However, it is not meaningful to compare the standardized structural coefficients across contexts (.32 versus .35) (see Table 6.8 and 6.9).

Table 6.8: Structural coefficients for the offline context for base model

Structural relationships offline context N=564	Unstandardized structural coefficient	Standardized structural coefficient	*t*-value	Hypothesis testing
Antecedents of Perceived value (R^2=.621)				
H2a: Service quality → Perceived value	.26	.29	4.71	Supported
H3a: Merchandise quality → Perceived value	-.02	-.03	-.53	Not supported
H4: Price → Perceived value	-.46	-.61	-7.56	Supported
Antecedents of Purchase intentions (R^2=.374)				
H1: Perceived value → Intentions	.11	.07	1.33	Not supported
H5: Perceived risk → Intentions	-.17	-.10	-2.07	Supported
H6: Time/effort costs → Intentions	-.18	-.20	-4.25	Supported
H7: Enjoyment → Intentions	.35	.32	5.76	Supported
H2b: Service quality → Intentions	.13	.10	1.26	Not supported
H3b: Merchandise quality → Intentions	.15	.12	2.14	Supported
Antecedent of Perceived risk (R^2=.133)				
H2c: Service quality → Perceived risk	-.29	-.37	-5.18	Supported
Antecedents of Enjoyment (R^2=.366)				
Service quality → Enjoyment	.50	.41	6.98	-
Merchandise quality → Enjoyment	.31	.28	5.24	-
Antecedent of Time/effort costs (R^2=.058)				
Merchandise quality → Time/effort costs	-.33	-.24	-4.55	-

Note: Based on one-tailed tests, *t*-values greater than 1.65 are significant at $p<.05$; *t*-values greater than 2.33 are significant at $p<.01$.

Table 6.9: Structural coefficients for the online context for base model

Structural relationships online context N=573	Unstandardized structural coefficient	Standardized structural coefficient	*t*-value	Hypothesis testing
Antecedents of Perceived value (R^2=.539)				
H2a: Service quality → Perceived value	.16	.18	2.86	Supported
H3a: Merchandise quality → Perceived value	.07	.10	1.96	Supported
H4: Price → Perceived value	-.46	-.55	-6.76	Supported
Antecedents of Purchase intentions (R^2=.565)				
H1: Perceived value → Intentions	.07	.03	.72	Not supported
H5: Perceived risk → Intentions	-.14	-.09	-2.33	Supported
H6: Time/effort costs → Intentions	-.30	-.19	-4.38	Supported
H7: Enjoyment → Intentions	.48	.35	7.49	Supported
H2b: Service quality → Intentions	.48	.24	3.83	Supported
H3b: Merchandise quality → Intentions	.15	.10	1.89	Supported
Antecedent of Perceived risk (R^2=.225)				
H2c: Service quality → Perceived risk	-.59	-.48	-7.83	Supported
Antecedents of Enjoyment (R^2=.381)				
Service quality → Enjoyment	.66	.47	8.22	-
Merchandise quality → Enjoyment	.22	.21	4.06	-
Antecedent of Time/effort costs (R^2=.341)				
Merchandise quality → Time/effort costs	-.55	-.58	-11.28	-

Note: Based on one-tailed tests, *t*-values greater than 1.65 are significant at $p<.05$; *t*-values greater than 2.33 are significant at $p<.01$.

The hypotheses were tested through analyzing the *t*-values at a significance level of .05. Of the ten proposed hypotheses, three were not supported by the data in the offline context, whereas only one hypothesis was not supported in the online context. In both contexts, perceived value did not influence purchase intentions. In the offline context, merchandise quality did not affect perceived value, and service quality did not affect purchase intentions. Furthermore, evidence was found for three additional relationships in each context.

Offline perceived value (i.e. value for money) was determined by price (-.46) and service quality (.26), but not by merchandise quality. Online perceived value was determined by

price (-.46), service quality (.16), and merchandise quality (.07). The small effect of merchandise quality on perceived value can be explained as consumers often consider what they receive in exchange for the price they pay. In their eyes, a well-composed assortment may not add value, as the only tangible aspect they have after spending money is the book of interest. Another explanation is that prior studies (e.g. Baker et al. 2002) found a significant relationship between merchandise quality and perceived value, because they investigated retailers that offered more differentiated products and assortments. In these studies, retailers can increase perceived value by offering high-quality workmanship and/or well-composed assortments (Baker et al. 2002). In both contexts, price was the strongest predictor of value for money, followed by service quality and merchandise quality. As such, the construction of offline perceived value appeared to be very similar to that of online perceived value.

In both contexts merchandise quality and the shopping experience costs and benefits (time/effort costs, risk and enjoyment) directly impacted purchase intentions. In the online context, service quality also predicted purchase intentions. Surprisingly, in both contexts, value for money did not alter consumers' purchase intentions to buy through a particular channel[21]. It seems that the value for money consumers receive in the online and offline context does not alter their intentions to use a channel for purchasing[22]. The shopping experience costs and benefits, on the other hand, strongly affected the intentions to shop through the website or store. Shopping enjoyment ($\beta_{offline}$=.35/β_{online}=.48) and time and effort costs ($\beta_{offline}$=-.18/β_{online}=-.30) had a major influence on whether someone intends to buy online or offline. Risk perceptions, being part of the shopping experience, also had a

[21] Mediation tests were performed according to the procedure outlined by Baron and Kenny (1986). The results showed that perceived value considerably mediated the relationships of merchandise quality, service quality and price with intentions in each context. The effect of price on offline purchase intentions was *fully* mediated by perceived value. These results confirmed that perceived value is an important mediator of service quality, merchandise quality and price with purchase intentions.

[22] The models were also tested without the direct relationships of merchandise quality and service quality on purchase intentions. When omitting these relationships, the relationship between perceived value and purchase intentions became significant in both contexts. Still perceived value –based on the standardized total effects– had the weakest impact on intentions. Next, as the models stand in a nested sequence, the χ^2 difference was tested in each context. The addition of the two relationships resulted in a significant decrease in chi square with 2 *df* ($p<.05$). Hence, the current model is preferred.

significant direct effect on purchase intentions in both contexts. However, the effect of it was not as substantial as time/effort and enjoyment. In fact, when looking at the standardized structural coefficients, risk had half the effect of time/effort costs (-.10 versus -.20 in the offline context, -.09 versus -.19 in the online context) and less than a third of the effect of enjoyment (-.10 versus .32 in the offline context, -.09 versus .35 in the online context). The limited role of risk can be explained due to the relative low risk involved in buying books (cf. Chen and Dubinsky 2003).

In order to better understand the total influence of the exogenous factors on the endogenous factor purchase intentions, both direct and indirect effects were investigated. Indirect effects represent the influence of the exogenous factors on an endogenous factor as mediated by one or more intervening factors; they are derived by multiplying the unstandardized parameter estimates of the intervening factors[23] (Diamantopoulos and Siguaw 2000). Table 6.10 shows the total effects (direct and indirect effects) of the predictors of online and offline purchase intentions.

A comparison of the total unstandardized effects for the online and offline context shows that the weights of most coefficients are rather similar –except for service quality, time/effort costs, and enjoyment– indicating that corresponding criteria were used in explaining online and offline intentions. Based on the unstandardized effects, service quality[24] appeared to have a stronger impact on intentions in the online context (.89) than in the offline context (.38). Time and effort costs have a stronger unstandardized effect in the online context (-.30) relative to the offline context (-.18). Finally, it appeared that enjoyment played a somewhat stronger role in the online context (.48) than in the offline context (.35). Thus, in contrary to the expectations, enjoyment seemed to more strongly impact consumers' intentions to shop online. In both contexts, service quality had a strong

[23] Indirect effects that are based on the multiplicative computation of at least one insignificant path need to be interpreted with caution because they can be misleading, as the confidence intervals include zero and possibly estimates of the opposite sign (Howell 1987). The indirect effects that were mediated through perceived value therefore need to be interpreted with caution in both contexts.

[24] The very strong effect of service quality on purchase intentions in the online context can be explained as one of the items refers to the aspect of reliability/fulfillment (i.e. keeping promises), which plays a dominant role in the online context (Parasuraman and Grewal 2000; Parasuraman et al. 2005; Wolfinbarger and Gilly 2003).

indirect effect on purchase intentions by altering perceptions of enjoyment and risk. For the offline context, the indirect effect of service quality on purchase intentions (.25) was even greater than its direct effect (.13). Merchandise quality also demonstrated a strong indirect effect in each context, albeit it was less strong than that of service quality.

Table 6.10: Total (standardized) effects on purchase intentions for base model

	Offline context N=564				Online context N=573			
Total effects on purchase intentions	Total effects	Direct effect[a]	Indirect effect	Total standar-dized effect[b]	Total effects	Direct effect[a]	Indirect effect	Total standar-dized effect[b]
Service quality	.38	.13[n.s.]	.25	.29 (2)	.89	.48	.41	.46 (1)
Merchandise quality	.31	.15	.17	.26 (3)	.43	.15	.28	.29 (3)
Monetary price	-.05	-	-.05	-.05 (7)	-.03	-	-.03	-.02 (7)
Perceived risk	-.17	-.17	-	-.10 (5)	-.14	-.14	-	-.09 (5)
Time/effort costs	-.18	-.18	-	-.20 (4)	-.30	-.30	-	-.19 (4)
Enjoyment	.35	.35	-	.32 (1)	.48	.48	-	.35 (2)
Perceived value	.11	.11[n.s.]	-	.07 (6)	.07	.07[n.s.]	-	.03 (6)

Notes:

a. N.s. represents coefficients of direct effects that are not significant from zero at .05 based on one-tailed tests. Figures in bold represent construct reliabilities, which were calculated based on the formula provided by Hair et al. (1998, p. 624).

b. Figures between brackets indicate the ranking of each factor in explaining the endogenous latent variable.

Compared with the unstandardized effects, standardized effects are better capable of representing the relative contribution of the predictors in explaining purchase intentions within each context. Offline purchase intentions were most strongly affected by enjoyment (.32), service quality (.29), merchandise quality (.26), and time/effort costs (-.20). Similar results show up for the online context; online purchase intentions were most strongly influenced by service quality (.46), enjoyment (.35), merchandise quality (.29), and time/effort costs (-.19).

A central finding is that service quality, merchandise quality and the shopping experience costs and benefits largely explain intentions to use a channel for buying books. It seems that consumers take into account the same variables but attribute different scores to their channels. Most noteworthy is the strong impact of enjoyment on purchase intentions in both contexts. Although past literature indicates that shopping costs (e.g. time/effort costs, perceived risk) more strongly impact consumers (Babin and Darden 1996; Baker et al. 2002; Sweeney et al. 1999), this study finds evidence that shopping enjoyment at least equally affects consumers' purchase intentions. This finding underlines the importance of the hedonic aspects of shopping (Arnold and Reynolds 2003; Duman 2002; Hirschman and Holbrook 1982; Wakefield and Baker 1985).

6.6.2 Assessment of measurement invariance

The prior section provided insights into the effects of the exogenous factors on the endogenous factors. This section addresses the preparations to formally test the relative strength of specific relationships, namely the assessment of measurement invariance.

Childers et al. (2001) argued that disagreement exists among structural equation modeling experts on the necessity of assessing measurement invariance as a prerequisite to the comparison of structural parameters across samples (cf. Bollen 1989; Byrne, Shavelson and Muthén. 1989; Hayduk 1996; Horn and McArdle 1992; Jöreskog and Sörbom 1971; Little 1997). Contrastingly, Vandenberg and Lance (2000) are very clear that the establishment of measurement invariance across groups is a logical prerequisite to conducting substantive cross-group or cross-context comparisons, such as tests of the invariance of structural parameter estimates. Following their line of reasoning, analyses of the differences between the structural relationships can only be meaningful, when the items measure the same thing and to the same degree in both contexts. As Vandenberg and Lance (2000, p. 40) elegantly put it "Comparisons of apples to apples are meaningful. Comparisons of sandwiches to sand wedges are not." Although the items in this study are visually measuring the same, respondents may attribute other values or respond differently for the online versus offline context. For instance, respondents could relate other service elements to service quality online versus offline, or may answer with more extreme answers. In effect, this may cause that comparisons are not justifiable. Next, it can be true that online buyers use different

conceptual frames of reference and attach different meanings to constructs than offline buyers (cf. Cheung and Rensvold 2000; Riordan and Vandenberg 1994). The essential question therefore is "to what extent are manifest variables' (i.e. Xs') measurement properties transportable or generalizable across populations [or contexts]?" (Vandenberg and Lance 2000, p.8).

According to Steenkamp and Baumgartner (1998) multiple group confirmatory factor analysis represents the most powerful and versatile approach to testing for measurement invariance. If tested, measurement invariance generally proceeds through testing multiple increasingly restrictive stages. After an extensive review of the measurement invariance literature, Vandenberg and Lance (2000) proposed a guideline with a more detailed list of increasingly restrictive stages (see Table 6.11). The required level of measurement invariance is dependent on the goal of the study. For example, if researchers want to compare latent means across contexts, configural, metric and scalar invariance should be established before comparisons can be meaningful (Hong, Malik and Lee 2003; Meredith 1993; Steenkamp and Baumgartner 1998). Childers et al. (2001) argued that the invariance of structural relationships is generally tested through analyzing the (1) invariance of the hypothetical pattern (configural invariance), (2) invariance of factor loadings (metric invariance), and (3) invariance of disturbance variances and perhaps, covariances. Next they argued, along with many authors, that the third assessment is too restrictive (Byrne 2001; Byrne et al. 1989; Childers et al. 2001; Horn and McArdle 1992; MacCullum et al. 1994; Widaman and Reise 1997). This study complies with the argument that configural and metric invariance should be established for assessing the invariance of structural relationships. Accordingly, this study first tests whether the measurement model is equivalent for the offline and online context.

Table 6.11: Measurement invariance tests (Vandenberg and Lance 2000)

	Null hypothesis	Explanation
1. Omnibus test	$\Sigma^g = \Sigma^{g'}$	A test of the null hypothesis of invariant covariance matrices
2. Configural invariance test		Test of the null hypothesis that the a priori pattern of free and fixed factor loadings imposed on the measures components (e.g. items) is equivalent across groups.
3. Metric invariance test	$\Lambda_k^g = \Lambda_k^{g'}$	A test of the null hypothesis that the regression slopes linking the manifest items to the underlying constructs are invariant across groups.
4. Scalar invariance test	$\tau_k^g = \tau_k^{g'}$	A test of the null hypothesis that intercepts of like items' regressions on the latent variable(s) are invariant across groups
5. Uniqueness invariance test	$\Theta^g{}_{\delta k} = \Theta^{g'}{}_{\delta k}$	A test of the null hypothesis that like items' unique variances are invariant across groups
6. Factor-variance invariance test	$\Phi_j^g = \Phi_j^{g'}$	A test of the null hypothesis that factor variances are invariant across groups
7. Factor-covariance invariance test	**e.g.** $\Phi^g{}_{21} = \Phi^{g'}{}_{21}$	A test of the null hypothesis that factor covariances are invariant across groups.
8. Factor-means invariance test	$\kappa^g = \kappa^{g'}$	A test of the null hypothesis of invariant factor means across groups

Omnibus test. Vandenberg and Lance (2000) propose to first conduct an omnibus test to check whether there is overall measurement invariance across groups. Mostly this is performed by testing the equality of the groups' covariance matrixes. Failure to reject the null hypothesis is commonly interpreted as a demonstration of overall measurement invariance across groups. Then, no further invariance tests are required. If, however, the null hypothesis, stating that the covariance matrices are invariant, is rejected, further analyses are required in order to identify the source of nonequivalence (Byrne 2001, p. 126). The omnibus test showed that the covariance matrices were not equivalent by constraining the factor loadings, variances and covariances to be equal ($\Delta\chi^2$ 497.23 with 68 *df*, $p<.001$).

Configural invariance test. A test of the null hypothesis that the base model structure (i.e. the pattern of fixed and nonfixed parameters) is invariant across groups (Hong et al. 2003). According to Steenkamp and Baumgartner (2000) configural invariance is supported when the specified model with zero loadings on nontarget factors fits the data well in all groups, all salient factor loadings are significantly and substantially different from zero, and the correlations between the latent factors are significantly below unity. This test must be established in order for subsequent tests to be meaningful (Hong et al. 2003; Steenkamp and Baumgartner 2000; Vandenberg and Lance 2000). The stacked model has 308 degrees of freedom (154 degrees of freedom for each baseline model) and showed reasonable fit indices (χ^2/df= 2.70, GFI=.93, CFI=.96, NNFI=.95, RMSEA=.039). The results indicate that full configural invariance was established.

Metric invariance test. A test of the null hypothesis that factor loadings for like items are invariant across groups (Horn and McArdle 1992). This test is needed to ensure that different groups respond to the items in the same way. This test is performed by constraining the factor loading of like items to be equal across groups (Hong et al. 2003; Vandenberg and Lance 2000). This is a stronger test than configural invariance because –in addition to specifying an invariant factor pattern– the loadings of the like items within that pattern are now constrained to be equal. Factor loadings are the regression slopes relating the X_{jk} to their corresponding latent variables, and consequently represent the expected change in the observed score on the item per unit change on the latent variable (Vandenberg and Lance 2000). The metric invariance test showed that the model fitted well with the hypothesized model (χ^2/df= 2.69, GFI=.93, CFI=.96, NNFI=.95, RMSEA=.038). However, the resulting χ^2 difference test with 12 degrees of freedom appeared significant ($p<.002$). Due to the fact that full metric invariance is difficult to establish, some researchers propose relaxing it by establishing partial metric invariance (Byrne 2001; Byrne et al. 1989). They suggest that if the noninvariant items constitute a small proportion of the model, cross-group comparisons can still be made because the noninvariant items will not affect the comparisons to any meaningful degree. In line with this reasoning, it was investigated which factor loadings were not invariant (see Appendix III). Two out of twelve estimated factor loadings (SQ4 and Int3) appeared to be nonequivalent across contexts. Consequently, the invariance tests testing the invariance of the structural

relationships were performed in which two factor loadings were set free (partial metric invariance).

Based on the statistical analyses, it can be concluded that configural and partial metric invariance were established. This was deemed sufficient for testing for differences in structural coefficients between the online and offline context. Additionally, metric invariance was also investigated from a practical approach. Due to chi square's extreme sensitivity to sample size and model complexity other authors also suggest a more practical approach by investigating the worsening of the fit indices by constraining parameters to be equal across contexts (Byrne 2001; Childers et al. 2001; Little 1997; MacCallum et al. 1994). Little (1997) proposed that the equality of factor loadings is upheld when the NNFI decreases less than .05 after imposing equality constraints on all factor loadings. This was the case, as the NNFI did not drop more than .01 after imposing the equality constraints.

A. Testing the relative importance of criteria in the online and offline context

After establishing configural and partial metric invariance[25], the hypotheses regarding the strength of structural relationships were tested (Hypotheses 12-15). To be more precise, it was tested whether time/effort costs, perceived risk and merchandise quality had a stronger effect on purchase intentions, and whether enjoyment had a less pronounced effect in the online context. First, it was tested whether all structural path coefficients were invariant across contexts. The chi-square difference test with 13 degrees of freedom appeared to be significant ($p<.001$), indicating that the structural path coefficients were not invariant across contexts. Next, to identify the source of nonequivalence, each separate relationship was constrained and set to be free (Byrne 2001). The difference in chi square with 1 degree of freedom was used to investigate whether the strength of relationship differed online versus offline (cf. Baker et al. 2002; Childers et al. 2001; Einwiller 2003). None of the four hypotheses were supported (see Table 6.12). The strengths of the relationships between time/effort costs and purchase intentions were not significantly different across contexts ($\beta_{offline}=-.18/\beta_{online}=-.30$, $p=.183$). Next, there was no significant difference between the strength of relationships between enjoyment and purchase intentions

[25] After full configural invariance was established, ten out of twelve items appeared invariant across the online and offline context (see Appendix III). The NNFI dropped less than .01 after imposing the equality constraints on the measurement items.

($\beta_{offline}$=.35/β_{online}=.48, p=.147). Then, the effects of risk on purchase intentions were also invariant across contexts ($\beta_{offline}$=-.17/β_{online}=-.14, p=.766), indicating risk plays an equal role in affecting purchase intentions across contexts. Finally, merchandise quality did not have a stronger direct effect on purchase intentions in the online context ($\beta_{offline}$=.15/β_{online}=.15, p=.975). This latter insignificant finding can be explained as the same bookseller was considered, which had similar assortments online and offline. Consumers were therefore less likely to be motivated to shop online, because of superior online assortments. In sum, it can be concluded that the relationships between the shopping experience costs/benefits and merchandise quality with intentions were similar across contexts. The only structural relationships that significantly differed between contexts were (1) merchandise quality → time/effort costs, and (2) service quality → perceived risk. A closer investigation of the coefficients (see Table 6.12) shows that in the online context, service quality more strongly reduced risk than in the offline context. Additionally, merchandise quality more strongly reduced time/effort costs in the online context. It seems that improvements in merchandise quality in the online context lead to major time/effort savings. Although the difference in the strength of service quality on intentions appeared large ($\beta_{offline}$=.13/β_{online}=.48), the difference was just outside the level of significance (p=.053) (see Appendix III).

Table 6.12: Tests of invariant structural relationships offline context versus online context

	Hypothesis	Structural coefficient offline context[a,b]	Structural coefficient online context[a,b]	P-value	Hypothesis Testing
Time/effort costs → Intentions	H12: Stronger in online context	-.18	-.30	.180	Not supported
Enjoyment → Intentions	H13: Stronger in offline context	.35	.48	.231	Not supported
Perceived risk → Intentions	H14: Stronger in online context	-.17	-.14	.773	Not supported
Merchandise quality → Intentions	H15: Stronger in online context	.15	.15	.975	Not supported
Service quality → Perceived risk	-	-.26	-.59	**.000**	-
Merchandise quality → Time/effort costs	-	-.30	-.56	**.000**	-

Notes:

a. The shown unstandardized structural coefficients marginally differ from those in Table 6.8 and 6.9 due to the equality constraints of the factor loadings.
b. N.s. represents coefficients that are not significant from zero at a .05 significance level based on one-tailed tests.

B. Testing the moderating effect of prior online shopping experience in the online context

It is assumed that the strength of relationships in the online context can be attenuated or strengthened through the level of prior online shopping experience[26]. Based on Chapter 4,

[26] It was also tested whether online and offline buyers differed in terms of their structural relationships regarding the offline context. Of the thirteen structural relationships, only one differed between offline and online buyers; the relationship between service quality and

three hypotheses were related to the moderating effect of prior online shopping experience (H17-H19). To test for the moderating effect of prior experience, this study discerned between offline buyers (those without direct experience with shopping through the website) and online buyers (those with direct experience with shopping through the website).

The two subgroups were evaluated in terms of their unstandardized structural coefficients. The same procedure was used to assess the differences in the strength of the relationships between online and offline buyers. First, the baseline models were established. The two separate models had acceptable fit indices for the offline and online buyers, respectively (χ^2/df= 2.04/1.89, GFI=.93/.89, CFI=.96/.94, RMSEA=.052/.068). Next, the omnibus test showed that the covariance matrices were not equivalent by constraining the factor loadings, variances and covariances to be equal ($\Delta\chi^2$ 423.38 with 68 *df*, p<.001). Subsequently, configural invariance was tested. The stacked model with 308 degrees of freedom showed reasonable fit indices (χ^2/df= 2.09, GFI=.90, CFI=.94, NNFI=.93, RMSEA=.043), indicating that full configural invariance was established. Next, it was investigated whether the factor loadings were invariant across online and offline buyers. The metric invariance test showed that the model fitted well with the hypothesized model (χ^2/df= 2.12, GFI=.90, CFI=.94, NNFI=.92, RMSEA=.044). However, the χ^2 difference test appeared significant ($\Delta\ \chi^2$=35.02 with 12 *df*, p<.001). A subsequent investigation showed that three out of twelve estimated factor loadings (Time2, PV1, and Int1) were nonequivalent across contexts. From a practical perspective, it was found that the factor loadings were equal, based on the criterion that the NNFI should not decrease more than .05 when full equality constraints are imposed (Little 1997). This was the case, as the NNFI marginally dropped (Δ NNFI=.002) after imposing the equality constraints. Consequently, the structural invariance tests were performed after configural and partial metric invariance were established. Table 6.13 shows the results of the hypotheses and two nonhypothesized significant differences.

enjoyment was stronger for offline buyers (.66) than for online buyers (.33). Pooling the data seemed justifiable, as it did not substantially alter the relationships.

Table 6.13: Tests of invariant structural relationships offline buyers versus online buyers

	Hypothesis	Structural coefficient offline buyers	Structural coefficient online buyers	P-value	Hypothesis testing
Perceived risk → Intentions	H17: Attenuated by prior online shopping experience	-.17	.02 n.s.	.136	Not supported
Time/effort costs → Intentions	H18: Strengthened by prior online shopping experience	-.14	-.27	.181	Not supported
Enjoyment → Intentions	H19: Attenuated by prior online shopping experience	.56	.16	**.000**	Supported
Service quality → Enjoyment	-	.83	.43	**.027**	-
Service quality → Intentions	-	.50	-.14 n.s.	**.001**	-

Note: N.s. represents unstandardized coefficients that are not significant from zero at a .05 significance level based on one-tailed tests.

The results suggest that online buyers rely less strongly on risk and more strongly on time/effort savings than offline buyers. However, these differences were not significant. Therefore hypotheses H17-H18 could not be supported, although the differences were in the right direction. Hypothesis 19, stating that online buyers are less affected by enjoyment than offline buyers, was confirmed. Based on the item means (see section 6.3), it appears that offline buyers strongly rely on enjoyment because of the lack of enjoyment in the

online context. A relatively small improvement in enjoyment strongly increases the offline buyers' online purchase intentions[27].

The two nonhypothesized significant differences can be explained as follows: for offline buyers, service quality is very important as it entails the aspect of reliability/fulfillment. As a result, service quality has a stronger impact on enjoyment and intentions for offline buyers than for online buyers. If e-tailers succeed in improving offline buyers' online service quality perceptions, offline buyers will perceive the online shopping process to be more enjoyable, and will have higher online purchase intentions.

6.7 Stage 5: Multiple group CFA for extended model

The effect of reputation, informativeness and ease of use was investigated in the extended model. The goal of this extension was to ensure no important factors were left out for the online context. Next, these additional factors also provide insights into how the predictors of perceived value and purchase intentions are constructed.

Appendix IV shows the item means, standard deviations and mean differences between the online and offline context for the two groups of buyers. Offline buyers found the offline channel easier to use, more informative, and evaluated the store's reputation much higher than the website's reputation. Online buyers generally considered the online channel to be easier to use, but rated the reputation of the store higher than that of the website. They appeared indifferent towards the channels' capability to provide relevant information.

Separate exploratory factor analyses were run and found the three expected underlying factors. The KMO measure was .833 for the offline context, and .886 for the online context, supporting factor analysis of the data (Hair et al. 1998). The three factors explained 60.9% in the offline context and 65.7% in the online context. Then, in each context a CFA was run based on the 11 latent factors. The offline sample consisted of 539 buyers (372 offline buyers, 167 online buyers); the online sample consisted of 502 buyers

[27] This partly explains why enjoyment had a substantial effect (β=.48) on purchase intentions in the online context (see Hypothesis 13).

(330 offline buyers, 172 online buyers). The measurement model fitted the data well for the offline and online context (χ^2/df=2.05/2.22, GFI=.95/.94, CFI=.96/.95, NNFI=.95/, RMSEA=.044/.050). The three constructs appeared reliable and possessed sufficient convergent validity (see Table 6.14). Reliabilities were above the recommended level of .60 (Bagozzi and Yi 1988), and the *t*-values were significant (p<.001). In addition, all constructs –except offline ease of use– met the minimum AVE level of .50. However, the discriminant validity of the constructs was questioned, when analyzing the correlations between the latent factors. Reputation was highly correlated with service quality ($\varrho_{offline}$=.84, ϱ_{online}=.86), and ease of use ($\varrho_{offline}$=.60, ϱ_{online}=.75). Next, ease of use was highly correlated with information relevancy ($\varrho_{offline}$=.62, ϱ_{online}=.74). This high degree of correlation could lead to multicollinearity problems. The VIFs were investigated in subsequent regression analyses with unweighted summated scores. The largest VIF value in the offline context was 2.45 in the offline context and 2.79 in the online context, indicating that there was some –but no severe– degree of multicollinearity (Parasuraman et al. 2005). One of the consequences of multicollinearity is that it makes determining the relative contribution of each independent factor more difficult because the effects of the independent factors are mixed or confounded (Hair et al. 1998). One of the remedies is to present the bivariate correlations between the independent and dependent factor in order to understand its relationship (Hair et al. 1998). Hence, this study tests the relationships, and –in case of insignificant relationships– shows the bivariate correlations to address the association between the two variables in isolation.

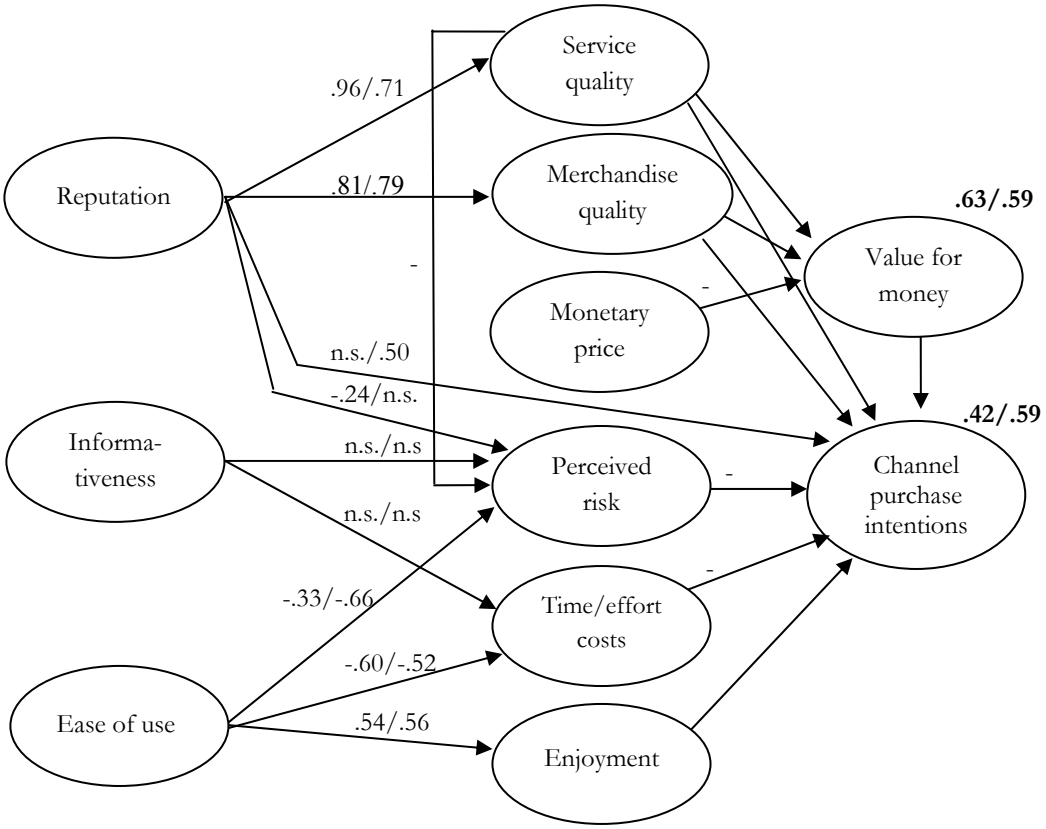
Table 6.14: Item loadings, construct reliabilities and average variance extracted

	Offline context N=539		Online context N=502	
	Standardized loading[a,bc]	**AVE[d]**	**Standardized loading[a,b,c]**	**AVE[d]**
Reputation	**.88**	*.70*	**.89**	*.74*
Rep1: trustworthiness	.78		.83	
Rep2: good reputation	.87 (20.75)		.88 (23.48)	
Rep3: reputable standing	.86 (20.46)		.85 (22.27)	
Informativeness	**.87**	*.68*	**.86**	*.67*
Inf1: in-depth information	.80		.73	
Inf2: relevant information	.92 (18.33)		.85 (18.01)	
Inf3: right information	.75 (21.20)		.87 (18.27)	
Ease of use	**.68**	*.42*	**.75**	*.50*
Ease1: access convenience[e]	.51		.61	
Ease3: search convenience	.65 (9.41)		.73 (12.36)	
Ease5: shopping convenience	.75 (9.84)		.77 (12.68)	

Notes:

a. Figures in bold represent construct reliabilities, which were calculated based on the formula provided by Hair et al. (1998, p. 624).
b. Figures between brackets represent *t*-values of the factor loadings. The first item of each construct was used as a reference item.
c. Based on one-tailed tests, *t*-values greater than 1.65 are significant at $p<.05$; *t*-values greater than 2.33 are significant at $p<.01$.
d. The average variance extracted (AVE) was calculated based on the formula provided by Fornell and Larcker (1981).
e. For the exact wording of the items, see Table 5.1.

Again, the baseline models were first established, prior to testing the invariance of the structural relationships. Figure 6.2 shows the results of the unstandardized coefficients of the three factors.



Fit indices	Offline context	Online context
χ^2	793.62	887.46
df	349	349
χ^2/df	2.27	2.54
GFI	.90	.89
AGFI	.88	.86
NNFI	.94	.93
CFI	.95	.94
RFI	.89	.89
SRMR	.058	.064
RMSEA	.049	.055

Figure 6.2: Coefficients for offline/online context for extended model

Notes: The unstandardized structural coefficients are displayed for the offline/online context. Figures in bold represent the percentage of explained variance in the endogenous variables. N.s. represents coefficients that are not significant from zero at a .05 significance level based on one-tailed tests.

The model fitted the data well in each context, although the fit indices were somewhat lower than in the base model. The modification indices suggested no additional relationships. Table 6.15 shows the unstandardized coefficients with their corresponding *t*-values. The hypotheses were tested through analyzing the *t*-values at a significance level of .05. Of the nine proposed hypotheses, three were not supported by the data in each context.

As expected, reputation strongly altered perceptions of service quality ($\beta_{offline}$=.96, β_{online}=.71) and merchandise quality ($\beta_{offline}$=.81, β_{online}=.79). In the offline context, reputation significantly reduced risk ($\beta_{offline}$=-.24), but it did not have a direct impact on purchase intentions. In the online context the opposite was found, reputation did not reduce risk[28], but it directly affected purchase intentions (β_{online}=.50). The reputation of the website thus had a strong direct influence on intentions, indicating the prominent role reputation plays online. In both contexts, information relevancy neither reduced risk perceptions nor time/effort costs. It seems that ease of use explained most of the variance in the endogenous factors, as the correlations suggest that informativeness is moderately correlated with time/effort costs ($\varrho_{offline}$=.-.27, ϱ_{online}=-.47) and with risk ($\varrho_{offline}$=-.23, ϱ_{online}=-.43). Ease of use had a strong influence on time/effort costs ($\beta_{offline}$=-.60, β_{online}= -.52), risk ($\beta_{offline}$=-.33, β_{online}=-.66), and enjoyment ($\beta_{offline}$=.54, β_{online}=.56), and thus strongly affected both utilitarian and hedonic aspects of shopping.

[28] Although this relationship was insignificant in the online context, the correlation between reputation and risk (-.52) indicated that they were significantly correlated.

Table 6.15: Unstandardized structural coefficients for additional factors

	Offline context N=539			Online context N=502		
Hypothesis	**Structural coefficient**	***t*-value**	**Hypothesis testing**	**Structural coefficient**	***t*-value**	**Hypothesis testing**
H9a: Reputation → Service quality	.96	15.48	Supported	.71	15.61	Supported
H9b: Reputation → Merchandise quality	.81	13.78	Supported	.79	14.41	Supported
H9c: Reputation → Perceived risk	-.24	-1.66	Supported	.04	.21	Not supported
H9d: Reputation → Intentions	.32	1.58	Not supported	.50	2.28	Supported
H10a: Informativeness → Time/effort costs	-.06	-.77	Not supported	.11	1.46	Not supported
H10b: Informativeness → Perceived risk	.05	.91	Not supported	.09	1.07	Not supported
H11a: Ease of use → Time/effort costs	-.60	4.20	Supported	-.52	6.42	Supported
H11b: Ease of use → Perceived risk	-.33	-3.28	Supported	-.66	-5.26	Supported
H11c: Ease of use → Enjoyment	.54	5.60	Supported	.52	6.98	Supported

Note: Based on one-tailed tests, unstandardized coefficients with *t*-values greater than 1.65 are significant at $p<.05$; *t*-values greater than 2.33 are significant at $p<.01$.

The unstandardized direct effects provide useful information, but to better understand the total influence of the additional factors on the endogenous factors, the total (standardized) effects were analyzed (see Table 6.16). The addition of the three factors altered some of the relationships in the base model. For example, the addition of reputation as a predictor of online purchase intentions made the relationships between merchandise and service quality with purchase intentions insignificant. Appendix V shows the differences in the strength of the coefficients due to the addition of the three factors.

Table 6.16: Total (standardized) effects for extended model

	Offline context N=539				Online context N=502			
Total effects on purchase intentions	Total effects	Direct Effects[a]	Indirect effects	Total standar-dized effect[b]	Total effects	Direct effects[a]	Indirect effects	Total standar-dized effect[b]
Service quality	.12	-.05[n.s.]	.17	.08 (6)	.23	.08[n.s.]	.15	.11 (5)
Merchandise quality	.19	.14	.05	.15 (4)	.14	.08[n.s.]	.06	.10 (6)
Monetary price	-.04	-	-.04	-.04 (9)	-.03	-	-.03	-.02 (9)
Enjoyment	.36	.36	-	.33 (2)	.54	.54	-	.37 (2)
Time/effort costs	-.15	-.15	-	-.15 (5)	-.20	-.20	-	-.12 (4)
Perceived risk	-.13	-.13	-	-.08 (7)	-.08	-.08[n.s.]	-	-.05 (7)
Perceived value	.10	.10[n.s.]	-	.07 (8)	.07	.07[n.s.]	-	.03 (8)
Reputation	.62	.32[n.s.]	.30	.39 (1)	.77	.50	.27	.44 (1)
Informativeness	.00	-	.00	.00 (10)	-.03	-	-.03	-.02 (10)
Ease of use	.32	-	.32	.20 (3)	.45	-	.45	.30 (3)
Total effects on perceived value								
Monetary price	-.44	-.44	-	-.62 (1)	-.49	-.49	-	-.56 (1)
Service quality	.31	.31	-	.34 (2)	.31	.31	-	.31 (2)
Merchandise quality	-.03	-.03[n.s.]	-	-.04 (4)	.01	.01[n.s.]	-	.03 (4)
Reputation	.27	-	.27	.26 (3)	.23	-	.23	.29 (3)
Total effects on perceived risk								
Service quality	.04	.04[n.s.]	-	.05 (4)	-.13	-.13[n.s.]	-	-.09 (2)
Reputation	-.20	-.24	.04	-.21 (2)	-.05	.04[n.s.]	-.09	-.05 (4)
Informativeness	.05	.05[n.s.]	-	.06 (3)	.09	.09[n.s.]	-	.08 (3)
Ease of use	-.33	-.33	-	-.33 (1)	-.66	-.66	-	-.66 (1)

Table 6.16: Total (standardized) effects for extended model (continued)

	Offline context N=539				Online context N=502			
Total effects on time/effort costs	Total effects	Direct Effects[a]	Indirect effects	Total standar-dized effect[b]	Total effects	Direct effects[a]	Indirect effects	Total standar-dized effect[b]
Merchandise quality	-.02	-.02[n.s.]	-	-.01 (3)	-.26	-.26	-	-.29 (2)
Reputation	.01	-	.01	.01 (4)	-.21	-	.21	.19 (3)
Informativeness	-.06	-.06[n.s.]	-	-.05 (2)	.11	.11[n.s.]	-	.10 (4)
Ease of use	-.60	-.60	-	-.37 (1)	-.52	-.52	-	-.55 (1)
Total effects on enjoyment								
Service quality	.31	.31	-	.31 (3)	.22	.22	-	.15 (2)
Merchandise quality	.15	.15	-	.13 (4)	.02	.02[n.s.]	-	.02 (4)
Reputation	.50	-	.50	.34 (2)	.17	-	.17	.15 (3)
Ease of use	.54	.54	-	.35 (1)	.56	.56	-	.54 (1)

Notes:

a. N.s. represents coefficients of direct effects that are not significant from zero at .05 based on one-tailed tests. Figures in bold represent construct reliabilities, which were calculated based on the formula provided by Hair et al. (1998, p. 624).

b. Figures between brackets indicate the ranking of each factor in explaining the endogenous latent variable.

Demonstrated by the total standardized effects, the additional factors ease of use and reputation had a strong impact on purchase intentions. Reputation –being highly a strong predictor of both service and merchandise quality– took over the effect of service and merchandise quality and had the most substantial effect on purchase intentions in the online and offline context. Ease of use, which strongly affected the shopping experience costs and benefits, had the third-largest impact on purchase intentions. Informativeness had a marginal effect on purchase intentions[29].

[29] This does not imply that providing relevant information is not important to online and offline booksellers. Informativeness was strongly correlated with ease of use (see above),

When investigating the influencers of perceived value, it appears that reputation did have an impact through altering perceptions of service quality and merchandise quality. However, the impact of reputation was less than the effect of price and service quality on value in each context.

Perceived risk was most strongly reduced by ease of use in both contexts. This effect seems to be particularly pronounced in the online context. Thus, online booksellers can effectively reduce risk by making the shopping process more convenient. Although past literature indicated that online risk is effectively reduced by increasing the reputation of the website (e.g. Einwiller 2003), this study finds evidence that making the shopping process more convenient is more appropriate. In the offline context, reputation did reduce risk, but the effect was not as substantial as that of ease of use.

Not surprisingly, time/effort costs were predominantly explained by ease of use. Customers who find the store/website easy to use can more easily and quickly obtain their desired book and save time and effort. Providing more relevant information to customers, however, did not lead to time/effort savings in either context.

Enjoyment was also largely determined by ease of use in each context. A more convenient shopping process is more likely to evoke positive affect (Childers et al. 2001). In the offline context, reputation had a substantive impact on enjoyment through affecting both service quality and merchandise quality. In the online context, enjoyment only seemed to stem from the ease of use offered by the website; its effect was more than three times as great as the second-strongest predictor.

Overall, the results suggest that ease of use and reputation significantly altered perceptions of perceived value and intentions, whereas informativeness did not. Reputation for a large part affected purchase intentions through altering merchandise and service quality[30],

---

and it had high levels of shared variance with ease of use. Informativeness did not explain a significant amount of unique variance in the endogenous variables.

[30] The total indirect effect of reputation on purchase intentions was based on a number of multiplicative computations with insignificant paths. The indirect effects derived from these insignificant paths were, however, relatively small compared to the indirect effects derived from significant paths.

whereas ease of use did this by strongly affecting the shopping experience costs and benefits. Next, based on the examination of the total effects, there is little evidence for major differences in the importance of criteria between channels. The total effect of ease of use appears to be somewhat more pronounced in the online context, but the difference is not substantial. Thus, although some authors argued that ease of use, information relevancy, and reputation play a more profound role in the online context; this study does not find clear evidence for this, as the factors play a similar role in each context. The next sections address the relative strength of the relationship of reputation on risk across contexts (Hypothesis 16) and the moderating effect of prior online shopping experience on the relationship between reputation and risk (Hypothesis 20).

A. Testing the relative importance of reputation on risk in the online and offline context

Hypothesis 16 stated that reputation would more strongly reduce risk in the online context due to the absence of intrinsic product attributes. After configural and partial metric invariance was established[31], the strength of the relationship did not differ between contexts ($p=.224$) (see Table 6.17). Based on this, there was no support for Hypothesis 16. The structural path coefficients of the 22 relationships, however, showed that the relationships were not invariant across the online and offline context ($p<.001$). For the sake of completeness, the relationships with regard to ease of use, information and reputation were investigated. Although it seemed that the relationship between reputation and intentions was stronger in the online context ($\beta_{offline}=.32$, $\beta_{online}=.50$), there was not a significant difference between the two contexts ($p=.627$). Three relationships differed between contexts. Ease of use had a stronger impact on risk and enjoyment in the online context than in the offline context. When consumers find the website easier to use, it strongly reduces their risk perceptions and also drastically increases the shopping enjoyment. The results also show that the store's reputation more strongly influences service quality than the website's reputation. A possible explanation for this is that consumers are more familiar with the store than with the website and that the store's reputation is strongly tied to service quality.

[31] After full configural invariance was established, thirteen out of eighteen items appeared invariant across the online and offline context. The NNFI dropped less than .01 after imposing equality constraints on the measurement items.

Table 6.17: Tests of invariant structural relationships offline versus online context

	Hypothesis	Structural coefficient offline context[a]	Structural coefficient online context[a]	P-value	Hypo-thesis testing
Reputation → Perceived risk	H16: Stronger in online context	-.24	.04[n.s.]	.224	Not supported
Ease of use → Perceived risk	-	-.33	-.66	**.000**	-
Ease of use → Enjoyment	-	.54	.84	**.016**	-
Reputation → Service quality	-	.96	.71	**.000**	-

Notes: N.s. represents unstandardized coefficients that are not significant from zero at .05. Unstandardized structural coefficients may differ from those in Table 6.15 due to the equality constraints of the factor loadings.

B. Testing the moderating effect of prior online experience in the online context

Hypothesis 20 argued that the relationship between reputation and risk would be attenuated by the level of prior online experience. After configural and partial metric invariance was established[32], the structural invariance of this relationship was tested. The results did not lend support for Hypothesis 20 (see Table 6.18). Again the invariance of the structural relationships with regard to ease of use, information and reputation were investigated. Only one relationship significantly differed between the two groups; online buyers relied more heavily on reputation of the website as indicator of service quality. Online buyers are more familiar with shopping through the website, and this may explain why the relationship between reputation and service quality is stronger for them than those that have lower levels of familiarity.

[32] After full configural invariance was established, seventeen out of eighteen items appeared invariant across online and offline buyers (see Appendix III). The NNFI dropped less than .01 after imposing equality constraints on the measurement items.

Table 6.18: Test of invariant structural relationships offline versus online buyers

	Hypothesis	Structural coefficient offline buyers	Structural coefficient online buyers	P-value	Hypo-thesis testing
Reputation → Perceived risk	H20: Attenuated by prior online shopping experience	.10$^{n.s.}$	.07$^{n.s.}$	.960	Not supported
Reputation → Service quality	-	.62	.89	**.010**	-

Note: N.s. represents unstandardized coefficients that are not significant from zero at .05.

6.8 Stage 6: Discussion of findings

Overall, the results supported the proposed conceptual model. Most proposed relationships in the base model were confirmed by the data. In both contexts, purchase intentions were predominantly defined by service quality, merchandise quality, enjoyment and time/effort costs. Service quality and merchandise quality had strong indirect effects in each context. Service quality affected enjoyment and risk, whereas merchandise quality altered perceptions of time/effort costs and enjoyment. The results also demonstrated that enjoyment played a significant role and should be incorporated in perceived value models. Surprisingly, perceived value –defined as the value for money consumers receive– did not alter purchase intentions in either context. It seems that altering the value for money consumers receive hardly stimulate them to use a particular channel. To motivate them to use a particular channel, it is better to focus on improving the four above-mentioned factors.

It was investigated whether the strength of relationships differed between contexts (i.e. channels) by using a multiple group confirmatory factor analysis. None of the proposed hypotheses regarding the differences in the strength of motivations (risk, time/effort costs, enjoyment and merchandise quality) could be confirmed. In search for nonequivalence, two relationships appeared to be different across contexts: in the online context service quality had a stronger impact on risk, and merchandise quality had a stronger effect on

time/effort costs. Service quality seemed to have a somewhat more pronounced effect in the online context, but overall the results suggest that the construction of perceived value and intentions appear very similar. Customers do not significantly differ in the weights they attribute to the factors, but rather attribute different scores to the performance of each channel on these factors.

It was also investigated whether the strength of relationships differed between online buyers and offline buyers, i.e. customers who had prior experience with buying books through the bookseller's website and those who had no prior direct experience with the website. Compared to the offline buyers, online buyers relied less on enjoyment, but the two groups relied equally on time/effort savings and risk. Based on a comparison of the item means pertaining to enjoyment, it appears that the *lack* of enjoyment inhibits offline buyers to shop online; offline buyers therefore are strongly affected by enjoyment.

Next, the extended model was tested. Three additional factors (ease of use, informativeness and reputation) were included to ensure that no important predictors were left out for the online context. The results indicated that ease of use and reputation had strong effects on purchase intentions in the online context, but also in the offline context. Ease of use strongly impacted the shopping experience costs and benefits, whereas reputation largely explained service and merchandise quality.

For the extended model it was hypothesized that reputation would more strongly reduce risk perceptions in the online context. However, no support was found. Finally, it was proposed that online buyers would rely less on reputation than offline buyers in the online context; again no support was found.