Durham Research Online

Deposited in DRO:
13 November 2017

Version of attached file:
Accepted Version

Peer-review status of attached file:
Peer-reviewed

Citation for published item:

Further information on publisher’s website:
https://doi.org/10.1108/ijrdm-11-2013-0208

Publisher’s copyright statement:
This article is © Emerald Group Publishing and permission has been granted for this version to appear here http://dro.dur.ac.uk/23485. Emerald does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Emerald Group Publishing Limited

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full DRO policy for further details.
Examining retail customer experience and the moderation effect of loyalty programmes

Zhibin Lin & Dag Bennett

(Accepted 11 April 2014)


Structured Abstract:

- **Purpose:** To examine the construct of retail customer experience (CE) and its links to satisfaction and loyalty; and to test whether loyalty programmes perform a moderating effect on those links.
- **Design/methodology/approach:** A variety of retail attributes are integrated to develop a holistic customer experience construct using formative measures, with four in-built, differentiated replication studies conducted in the supermarket and department store sectors in China.
- **Findings:** The empirical results confirm the model of customer experience’s impact on customer satisfaction and loyalty; but reveal that loyalty programmes perform an insignificant moderating role in enhancing the linkages in the model.
- **Research implications:** Further studies may examine whether our findings hold true for each individual loyalty programme. The article calls for more studies based on multiple, in-built, differentiated replication studies and measures to encourage publication of negative empirical results so as to ensure empirical generalization and self-correction in the literature.
- **Practical implications:** Retail managers should focus attention on the design and delivery of great customer experience, without placing great reliance on loyalty programmes. Both cognitive and emotional attributes of retailing services should be considered for managing a holistic customer experience.
- **Originality/value:** The article examines a model of CE with loyalty programme as a possible moderator; it uses formative measures of CE, multiple in-built replications
and reports negative empirical results, which are critical to the development of scientific progress in retail management research.

**Keywords:** Customer experience, Loyalty programme, Satisfaction, Loyalty.

**Article Classification:** Research paper
Examining retail customer experience and the moderation effect of loyalty programmes

Introduction

Creating superior customer experience (CE) is a central concern in retail management and most managers recognize that enhancing CE is important for customer satisfaction, loyalty, and ultimately the firm’s profitability (Kumar et al., 2013). In order to provide good, positive CE, it is important to understand the factors that create CE, how to measure it and how it affects customer satisfaction and loyalty in different service settings (Klaus and Maklan, 2012). However, despite on-going conceptual development of CE and growing attention to the construct, only a limited amount of research has attempted to measure CE and test its effects on customer satisfaction and loyalty (Lemke et al., 2011, Maklan and Klaus, 2011, Verhoef et al., 2009), and those that do (e.g. Bagdare and Jain (2013), use reflective rather than formative measures of the construct. This is a pity because the CE construct is more appropriately measured formatively because the causal relationship is between the indicators and latent construct, i.e. each indicator influences CE, rather than the other way round (Cenfetelli and Bassellier, 2009, Diamantopoulos and Winklhofer, 2001). Moreover, authors exploring the measurement of CE rarely link the construct to key customer outcomes such as satisfaction and loyalty, ignoring possible mediating or moderating variables (e.g. Bagdare and Jain, 2013).

For many retail managers one of the most popular tools used to enhance CE, satisfaction and loyalty is a loyalty programme (Gable et al., 2008). Firms in various industries around the world have adopted programmes that offer incentives, rewards and benefits for enhanced customer loyalty (Yi and Jeon, 2003). Such programmes are very popular because they are generally straightforward in concept (if not execution). And yet,
despite their popularity, research on loyalty programmes has so far failed to reach consensus on their effectiveness. In fact the results between studies are often inconsistent or even contradictory (Kopalle et al., 2012, Liu, 2007, Reinartz and Kumar, 2003), and very few studies have addressed the potential moderating effect of loyalty programme membership on the links between loyalty antecedents and outcomes (Bolton et al., 2000, Evanschitzky and Wunderlich, 2006, Lacey and Morgan, 2009, Walsh et al., 2008).

Most research in retailing (and marketing in general) consists of one-off studies that often have elegant modelling and sophisticated analyses but only succeed in generating speculative results. Such studies, based on single sets of data are rarely re-tested or even confirmed (Ehrenberg, 1990, Schmidt, 2009) and therefore little useful knowledge is accumulated. This has left marketing with a ‘desperate need for replications’ (Hunter, 2001, p. 149) and therefore a growing number of scholars now advocate the use of differentiated replication studies using many sets of data that with sufficient depth can be used to establish empirical generalizations (Ehrenberg, 1990, Evanschitzky et al., 2007, Lindsay and Ehrenberg, 1993).

Against this background, the major aims of this study are: (1) to develop context-specific measures of CE in two major retail sectors; supermarkets and department stores, by adopting formative indicators based on key attributes of retailing services; (2) to empirically test a model which links CE to positive customer outcomes, i.e. satisfaction and loyalty; (3) to empirically test whether loyalty programmes perform a moderating effect on those links in the CE-satisfaction-loyalty model. The setting for the study is the retail environment of major urban centres in China, where retailers have benefited over the past decades from the opportunity to observe and possibly emulate international competitors entering the Chinese market. In doing so, Chinese retailers have quickly established high levels of competence
and competitiveness. Retailing in China therefore is something of a natural experiment and it is interesting to observe that loyalty programmes have rapidly become part of the competitive offering and customer experience.

In the following sections, we first briefly review relevant extant research on customer experience and loyalty programmes from which we develop a research model. Then we present the empirical results addressing the hypotheses, and conclude with a discussion of the managerial implications, research limitations and suggestions for future research.
Conceptual framework

Conceptualisation of retail customer experience

Previous academic studies of the antecedents of satisfaction and loyalty have mostly focused on service quality, such as the well cited SERVQUAL model (Parasuraman et al., 1988). This view of service quality has been criticised for being rather transaction-specific (Voss et al., 2008), and its focus on customer cognition (Liljander and Strandvik, 1997). Nevertheless the notions of service marketing have informed the development of experiential marketing (Grove and Fisk, 1997). Drawing on environmental psychology theory, Bitner (1992) proposes the concept of ‘servicescape’, which is the ‘total configuration of environmental dimensions’, including both the physical environment and the service staff as essential components in creating service experiences (Berry et al., 2002, Booms and Bitner, 1981, Hoffman and Turley, 2002). In contrast to the then prevailing focus on customer cognition, Holbrook and Hirschman (1982) argue for a shift of focus on the emotional aspects of consumption experience e.g. the pursuit of fantasies, feelings, and fun. Pine and Gilmore (1998) further urge companies to manage a service as a dramatic experience, similar to a theatre production. Empirical studies tend to support the role of emotive aspects of consumer behaviour (e.g. Dennis et al., 2013, Kozinets et al., 2002). Nevertheless, as suggested by Zomerdijk and Voss (2010), there seems to be a bias in the experiential marketing literature towards high-end retail sectors, and it is an open question whether the ordinary and mundane events such as grocery shopping should focus on ‘fantasies’.

Recent conceptual development of CE recognises the importance of both customer cognition and emotion (Edvardsson, 2005, Jüttner et al., 2013, Schembri, 2006). CE develops throughout all touch points and encounters during the service delivery process, but transcends the touch points, service encounters and processes (Jüttner et al., 2013), which
include pre- and post-purchase episodes, as well as past service processes, and their influence on future experience formation (Meyer and Schwager, 2007, Tynan and McKechnie, 2009, Zomerdijk and Voss, 2010). Therefore, CE is conceptualised as a holistic construct that is not restricted to the service and the service experience (Klaus and Maklan, 2012), but encompasses every aspect of a firm’s offering including all service encounters and interactions both with the service provider and fellow customers (Hume et al., 2006, Meyer and Schwager, 2007, Zomerdijk and Voss, 2010), and may include emotional, physical, intellectual, or even spiritual components (Gentile et al., 2007).

As a holistic concept, CE is most useful when operationalized in concrete frameworks that fit within the specific context of service (Bagdare and Jain, 2013, Grewal et al., 2009, Lemke et al., 2011, Palmer, 2010, Payne et al., 2008, Verhoef et al., 2009, Voss et al., 2008). Service design is then about orchestrating cognitive and emotional stimuli for the customers' experience (Berry et al., 2002). Stimuli are the “clue” (Berry et al., 2002), “cue” (Pine and Gilmore, 1999) or “touchpoint” (Zomerdijk and Voss, 2010), which ideally trigger positive cognitive and emotional responses from the customer (Berry et al., 2002). In practical terms, retail services can be seen as a bundle of core and peripheral services, or ‘service package’ of tangible and intangible attributes (Karwan and Markland, 2006). Defining the constituent parts helps generate practical implications for improving customer experience-enhancing strategies (Goldstein et al., 2002). Accordingly, customer experience needs to be deconstructed into its component retail service attributes (Meyer and Schwager, 2007), so as to generate insights for service design, which could orchestrate the physical environment or atmospherics (Kotler, 1973), people (customers and employees), and service delivery process to help customers co-create their desired experiences (Teixeira et al., 2012). Verhoef et al. (2009) suggest that factors of retail customer experience should include the factors of social environment, service interface, retail atmosphere, assortment, price and promotions, and
loyalty programmes. Grewal et al. (2009) identified promotion, price, merchandise, supply chain and location as key drivers for delivery of a superior retailer customer experience. However, none of these models have been empirically tested.

Following the holistic conceptualization of CE, we model CE as a formative construct derived from a selection of retail factors (Cenfetelli and Bassellier, 2009). The formative construct has several advantages: first, on a theoretical level formative measurement provides the means to model a CE from a potentially disparate set of customer experience factors such as merchandise quality, price, store atmosphere; second, at a nomological level, formative measurement facilitates the study of the causes and effects of a construct by bringing the analysis of varied indicators to the level of a holistic, single construct, CE; and third, because the measurable indicators in formative measurement models are regarded as uncorrelated, they can still be individually evaluated on their specific contributions to the construct by assessing their path weights (Cenfetelli and Bassellier, 2009).

**Outcomes of customer experience**

Two key outcomes of CE that are of great interest to retail managers are satisfaction and loyalty (Burns and Neisner, 2006, Grewal et al., 2009, Lu and Seock, 2008, Martenson, 2007, Meyer and Schwager, 2007, Sivadas and Baker-Prewitt, 2000, Verhoef et al., 2009). Customer satisfaction can be defined as the culmination of a series of customer experiences (Meyer and Schwager, 2007) and is widely seen as an important link to future customer patronage and referral behaviour (Seiders et al., 2005, Figure 1), which in turn affect a firm’s market share and profitability (Kumar et al., 2013).

Two major theories that researchers draw upon to support this conceptual framework are the ‘attitude-behaviour paradigm’ or the ‘cognitive-affect-behaviour’ model and the
‘stimulus-organism-response’ model (Mehrabian and Russell, 1974). Following the ‘cognitive-affect-behaviour’ model, CE and satisfaction can be seen to affect constructs that derive from cognitive evaluation of the marketing offer, which further leads to loyalty intention and behaviour. Similarly, following the ‘stimulus-organism-response’ model (Mehrabian and Russell, 1974), environmental stimuli such as the experience factors influence an individual customer’s emotional state (organism, e.g. CE and satisfaction), which in turn affects approach or avoidance responses, i.e. to stay loyal to retailers or to defect to other retailers.

Retail managers often focus on enhancing customer satisfaction through customer-orientated marketing strategies, and research generally supports the idea that customer satisfaction is an antecedent of customer loyalty (Mittal and Kamakura, 2001). Satisfied customers are assumed to be more willing to be engaged in cross- and up-buying of a firm’s products or services (Li et al., 2005), and also express higher repurchase intention, and actual repurchase (Mittal and Kamakura, 2001). We take this line of research further with the conceptual model depicted in Figure 1. Operationally, we define customer experience as a formative construct consisting of a retailer’s offering at attribute level, including merchandise, promotion, price, atmosphere, location, etc. depending on the specific research context (A1 through An). As suggested by Kumar et al. (2013), these attributes encompass not only the physical and concrete aspects of the retailing service, but also intangible elements such as store atmosphere and retailing reputation. Using both cognitive-affect-behaviour and stimulus-organism-response components we propose that:

H1. Retail customer experience has a positive effect on overall customer satisfaction.

H2. Customer satisfaction has a positive effect on customer loyalty.
Moderation effect of loyalty programme

A loyalty programme is defined as an organized marketing activity that offers a firm’s customers additional incentives, rewards or benefits to entice them to be more loyal (Leenheer et al., 2007, Liu, 2007, Sharp and Sharp, 1997, Yi and Jeon, 2003). The steady adoption of such programmes reflects acceptance of the relationship marketing idea (Liu and Yang, 2009) which can be seen as a paradigm shift from transaction-based marketing to customer retention and relationship management (Reichheld, 1993). The underlying rationale for loyalty programmes is usually based on the idea that: a) the cost to retain a customer is far less than the cost of acquiring a new one; and b) the Pareto rule which suggests that 80% of revenue comes from 20% of customers (Reichheld, 1993). It therefore seems reasonable to invest in retaining the firm’s most valuable customers. Also, in comparison to short-term effects of sales promotions and advertising campaigns (Ehrenberg et al., 1994), loyalty programmes seem to have a longer-term orientation (Sharp and Sharp, 1997).

In practice however, the effect of loyalty programmes on customer loyalty is hard to pin down (Dowling and Uncles, 1997, Hu et al., 2010, Liu, 2007). Researchers have searched for effects on a variety of customer related performance measures such as purchase behaviour (Drèze and Nunes, 2011, Meyer-Waarden and Benavent, 2009); share-of-wallet (Leenheer et al., 2007); tenure or life duration (Liu, 2007, Meyer-Waarden and Benavent, 2009), brand image, brand preference, satisfaction and affective attitude (Demoulin and Zidda, 2009, Hu et al., 2010), commitment (Lacey and Morgan, 2009), and price tolerance (Cortiñas et al., 2008). Yet in spite of this mass of work, there is no clear consensus on the effectiveness of loyalty
programmes—the evidence suggests that loyalty programmes mostly have no effect or where discernible, only weak effects. The only upside is that the research has identified a number of factors that may influence the performance of loyalty schemes including: the loyalty programme design, customer characteristics, the firm’s core product or service offerings and its relative market position (Liu and Yang, 2009).

Another approach to loyalty programme research aims to assess effectiveness by comparing the attitudes or behaviours of members of loyalty programmes against non-members. The problem with this approach is that it is difficult to establish a causal link since loyal customers are more likely to join a programme in the first place (Leenheer et al., 2007). Researchers also argue that accrued loyalty points serve as switching barriers, and should ‘have a moderating effect and play a crucial role in winning customer loyalty’ (Chen and Wang, 2009), but, little work supports this view since prior studies mostly focus on the direct effects of programmes on customer loyalty, and few address the potential moderating effect of loyalty programme membership on links between loyalty antecedents and outcomes (Bolton et al., 2000, Evanschitzky and Wunderlich, 2006, Lacey and Morgan, 2009, Walsh et al., 2008). Even within these studies results are mixed. Bolton et al. (2000) reported that programme membership moderated the relationships: a) between customer re-patronage intentions and repeat patronage; and b) between customer search and repeat patronage. In addition, Evanschitzky and Wunderlich (2006) confirmed the significant moderation effect of programme membership on the relationship between loyalty intention and action. However, they also found that membership’s moderating effects on the links between cognitive-to-affective, and affective-to-conative dimensions of loyalty were not significant. Lacey and Morgan (2009) indicated that membership positively moderates the relationship between commitment and customer desires to increase their purchases and provide complaint feedback to the firm, but does not affect other links between commitment and four key
advocacy outcomes. Walsh et al. (2008) and separately, Lacey and Morgan (2009) reported that loyalty programme membership did not show any positive moderator effects. Thus the extant loyalty programme literature offers few insights into the effects of loyalty programme membership on customer loyalty. Investigating the moderating effect of loyalty programme membership is therefore important because it addresses the fundamental purpose of such programmes, i.e. enhancing customer loyalty (Chen and Wang, 2009). In this paper we therefore test the moderation effect of loyalty programmes on the linkages of CE – satisfaction – loyalty:

H3. Loyalty programme membership positively moderates the relationship between customer experience and overall satisfaction;

H4. Loyalty programme membership positively moderates the relationship between overall satisfaction and loyalty.


Research design and methods

Research context

China’s retail market was historically composed of many small and medium-sized retailers concentrated in the economically well-developed eastern provinces, particularly in major cities such as Beijing, Shanghai and Shenzhen. With economic growth over the past decades came retail sector growth that has continued to the present day despite global recessions and recent currency crises. Since 2000 demand for major global brands has surged and modern retail formats such as supermarkets, department stores have been well received by Chinese consumers, including the international department store chains such as Parkson, Golden Eagle, Watson, New World and Ito-Yokado and international grocery retailers such as Wal-Mart, Auchan, Carrefour, and Tesco. In response to this strong new competition, local Chinese retailers have been quick to imitate and emulate the latest marketing and management concepts and many local retailers have started to focus on creating an engaging customer experience. This study therefore fits into a modern retail context and has implications for all retailers in China, both international and local.

Sampling and data collection

This study was conducted with a store intercept method based on the procedure used by Babin and Darden (1995) and Grace and O'Cass (2004) to collect data directly from shoppers in a shopping environment for each of the four studies with modifications from study to study. The basic rationale of using store intercepts is the method’s efficiency of accessing a large number of consumers and its ability to produce high quality (Grace and O'Cass, 2004). In order to ensure internal reliability and enhance the robustness of the data we specified four differentiated studies using separate data-sets from different retail categories. Following Babin and Darden (1995), in Study 1 we recruited three graduate students trained in market
research who approached respondents who were shopping at different supermarkets in 3 major cities in China: Beijing, Shanghai, and Shenzhen. The students personally interviewed the participants and recorded responses to the questionnaire. The average time for each interview was 3 minutes. No incentive was provided to respondents. In total, 315 valid responses were collected (Beijing n=106; Shanghai n=107, Shenzhen n=102). The response rate was approximately 25% in each city.

To be differentiated from Study 1, increase the response rate and get better cooperation from potential respondents, we secured the collaboration of an independent supermarket located in a town near Xiamen, southeast China for Study 2. Store check-out assistants (who were trained by a local market research consultancy) invited every 10th customer who had completed their shopping to participate in the survey. When consent was obtained, the store assistant then had respondents complete the questionnaire at a table next to the customer service counter near the store entrance. The store manager and the researcher from the consultancy supervised and assisted the fieldwork by responding to any queries raised. No monetary incentive was provided to respondents. The resulting sample was 211, a response rate of approximately 60%. The store’s manager confirmed that the samples were consistent with the general demographic profile of their customers.

Study 3 and 4 used a procedure similar to that for Study 2. The two department stores were high-end luxury chains. Instead of employing check-out assistants as in Study 2, six customer service assistants who regularly conduct the company’s in-house customer satisfaction surveys approached customers on a random basis when they were browsing in the store and a small gift (value = 10 CNY, or about 1 GBP) was provided to respondents as an incentive (funded by the collaborating department stores). Study 3 generated 241 valid responses, and Study 4 generated 300. The response rates for both studies were about 70%.
Overall the response rates in Study 2-4 are comparable to those obtained by Han et al. (2008) in similar settings (the exception was study 1, which had a slightly lower response rate). All store managers confirmed that the demographic profile of the samples was representative of the store’s clientele. We differentiated the fieldwork of Study 3 from Study 4 by collaborating with a different department store located in a separate district in Beijing, and Study 4 was conducted 6 months after Study 3.

Measures

Since CE is relatively new construct and no established items were available, we followed the widely adopted process of item generation as proposed by Churchill Jr (1979). We first specified the meaning and domain of CE based on insights from the literature (e.g. Carpenter and Moore, 2006, Dennis et al., 2002, Theodoridis and Chatzipanagiotou, 2009). This results in an initial list of 20 key items. We then presented them for review by four panels of experts, which consisted of 3 marketing academics - in Study 1, and 5 retail managers from each of the collaborating retailers in Study 2-4. Expert panels helped to reduce and refine the items and ensured face/content validity. Experts were asked to rate each item as “very important,” “somewhat important,” or “not important”. Only items rated very or somewhat important by a majority of experts were retained, and the wording of each item was short, simple, and to the point. The number of items for each of the four studies thus is different depending on its respective panel’s recommendations (details shown in Table 1).

For overall customer satisfaction and customer loyalty, we adapted the model commonly used in the American Customer Satisfaction Index, the European Customer Satisfaction Index and in the literature (Coelho and Henseler, 2012, Zeithaml et al., 1996).

Specifically, for satisfaction we adopted two items:

SAT 1: *Shopping at XYZ has been an enjoyable experience.*
SAT 2: Overall, how satisfied are you with XYZ?

For customer loyalty, we adopted three items:

LOY1: *I would consider XYZ my first choice to do shopping.*

LOY2: *I would recommend XYZ.*

LOY3: *I would do more shopping at XYZ in the next few years.*

The exact terms and phrases were modified from study 1 to study 4 to suit the retail context, while the meaning remained the same across the studies. The scale was differentiated from one study to another.
Results

We used Partial Least Square (PLS) structural equation modelling for our data analysis. Using PLS in path modelling to establish causal relationships is advantageous because it requires no assumption of normal distributions and also allows for smaller sample size requirements (Hair et al., 2011). The software used in this study is SmartPLS 2.0 M3 (Ringle et al., 2005). Unlike covariance-based structural equation modelling, PLS path modelling does not optimize a unique global scalar function, and subsequently does not have global goodness-of-fit. There have been attempts to develop overall model fit indices in PLS, such as the goodness-of-fit index (GoF) and the relative goodness-of-fit index (GoF$_{rel}$), but recent methodological research suggests that these indices are not suitable for PLS model validation (Henseler and Sarstedt, 2013). We therefore follow the commonly adopted guidelines as set by Hair et al. (2011), using a two-step procedure to assess the adequacy of the model. First, we assess the measurement model with regard to reliability and validity. Second, we examine the structural model parameters and the explanatory power of the model. To test the significance of model estimates, we compute the t-statistics using 5000 bootstrap samples (Hair et al., 2011). In the results tables, for the ease of reading we report ‘t-value of 1.96 and above’ in term of ‘p-value of 0.05 or lower’, and ‘t-value of 2.58 above’ in the term of ‘p-value of 0.01 or lower’.

Measurement model

We treat items measuring customer experience as formative indicators. Following Hair et al.’s (2011) recommendation, we examined multicollinearity among the indicators, and each indicator’s weight (relative importance) and loading (absolute importance). As shown in Table 1, all indicators’ variance inflation factors (VIFs) were lower than 5, thus multicollinearity was not a concern. Results of the item weights and loadings indicate that not
all the item weights are significant, but all the factor loadings are, except in Study 2, where neither the weights nor loadings of the two items ‘service personnel’ and ‘store atmosphere’ are significant and hence were not included for further analysis.

The weighting of the experience factors varied across the studies. This is typically the case in other studies as well where the importance of individual factors that ‘forms’ the holistic positive CE tend to be different. In the supermarket sector, Study 1 results indicate that the important attributes are: merchandise quality, service personnel, store atmosphere, and value for money. Study 2 shows that the important CE attributes are: merchandise quality, value for money, choice of products, promotional activities and price. In the high end department store sector, to the surprise of our collaborating retail managers, Studies 3 and 4 indicate that ‘promotional activities’ and ‘choice of big brands’ are not significantly important factors. Study 3 shows that service personnel, store atmosphere, and retailer reputation are important CE attributes, whereas, in Study 4 merchandise quality, store exterior, store layout, and location were important CE attributes.

Overall satisfaction and loyalty are treated as reflective constructs. We evaluate the measurement model by examining item loadings, composite reliability, convergent validity, and discriminant validity: all item loadings are above the recommended 0.7 (Hair et al., 2011), and all factor loadings for the two constructs are significant. The composite reliability (CR) exceeds the recommended level of 0.7, and the average variance extracted (AVE) values are above the recommended level of 0.5 (Hair et al., 2011) with the exception the AVE of loyalty in Study 4, which is 0.49.

(Table 1 about here)
Results in Table 2 indicate that the square roots of the AVE (where applicable) exceed the construct’s correlations with the other factors, thus discriminant validity can be established (Fornell and Larcker, 1981).

(Table 2 about here)

**Structural Model**

We use two nested models to test the main effect hypotheses (H1-H2) and the moderation hypotheses (H3 & H4). Model 1 tests the main link from customer experience, overall satisfaction and loyalty, and includes the control variables, i.e. age, gender, education and income. Model 2 adds loyalty programme membership as an independent variable, and the interaction term of ‘customer experience × loyalty programme membership’ and ‘overall satisfaction × loyalty programme membership’ following a residual centring approach (Chin et al., 2003).

Table 3 illustrates the model estimation results of Study 1 with a breakdown of the samples collected from different cities in China, i.e. Beijing, Shanghai and Shenzhen and the total samples. The aggregate PLS path coefficients are statistically significant: $R^2$ values for endogenous latent variables ‘satisfaction’ and ‘loyalty’ are fairly good. As hypothesized, customer experience is positively related to overall satisfaction (H1 was supported), overall satisfaction is positively related to customer loyalty (H2 was supported). Model 2 shows that the interaction terms ‘customer experience × loyalty programme membership’ and ‘overall satisfaction × loyalty programme membership’ are not significant. The R-squares in Model 2 have not shown much increase over Model 1. Thus H3 and H4 cannot be supported.

(Table 3 about here)
Table 4 presents the results of the Studies 2-4. The patterns in Study 1 were replicated with the three additional studies.

(Table 4 about here)
Discussion and Conclusions

This article attempted to measure the construct of CE in several specific settings and test its links to satisfaction and loyalty using loyalty programme membership as a moderator on those links through four in-built replication studies. We make three important contributions to the retailing literature:

First, we provide construct measurements and empirical evidence supporting the advances in the conceptualization of customer experience in retail management (Grewal et al., 2009, Verhoef et al., 2009). Unlike most prior research which adopts reflective measures of CE, we model CE as a formative construct based on the multiple factors of a retail offering while evaluating each factor’s contributions in forming the construct through their relative path weights, which are unavailable if the construct is treated as a reflective one (Cenfetelli and Bassellier, 2009, Diamantopoulos and Winklhofer, 2001). Our empirical evidence based on the relative path weights of each attributes suggests that the importance of individual attributes that ‘forms’ the holistic positive CE vary between the different studies. The results of Study 1, which was conducted in three Tier 1 developed cities, indicate that both cognitive and emotional attributes are important, which were largely replicated in Study 4 in context of department stores in Beijing, whereas in Study 2, which was conducted in a Tier 2 city (Xiamen), the important CE attributes are related more to cognitive factors than emotional ones. In contrast, only the results of Study 3 (high end department stores in Beijing) suggest that the affective aspects are more important than cognitive ones, supporting the findings of Dennis et al. (2002) and Dennis et al. (2013). Therefore, we may conclude that factors determining a positive retail CE tend to be context-dependent, potentially influenced by the stages of economic development and the degree of competition in the market, as well as individual characteristics (Carpenter and Moore, 2006, Theodoridis and Chatzipanagiotou,
The evidence supports the concerns raised by Zomerdijk and Voss (2010) that the ‘fun, fantasy’ experiential marketing (Holbrook and Hirschman, 1982) might be more suitable for high-end retail sectors in developed markets, than for grocery retailers in a less developed market, where the traditional marketing emphasis on customer cognition still applies. Overall, our findings reveal a variety of key factors of CE as suggested by Grewal et al. (2009) and Verhoef et al. (2009), support the holistic conceptualisation of CE (Berry et al., 2002, Jüttner et al., 2013, Schmitt, 1999, Tyan and McKechnie, 2009), and confirm that the CE model which demonstrates clear linkages between CE and satisfaction and customer loyalty.

Second, we test possible moderation effects of loyalty programme membership which has been largely neglected in the prior literature. We first examined the direct effect of loyalty programme memberships, and observed some positive signs, although they are not prevalent across all the studies. In Study 1, the path coefficient from membership to satisfaction is 0.21 (p<0.01) for the Shanghai sample, and in Study 2, the path coefficient from membership to loyalty is 0.23 (p<0.01), which seem to indicate a significant and positive direct effect of loyalty scheme membership on satisfaction/loyalty. As we coded the variable membership as 0=non-member, 1=member, the positive and significant relationships simply suggest that members are more satisfied/loyal than non-members in those cases. As argued by Leenheer et al. (2007), the result can be attributed to a ‘self-selection’ effect, i.e. more satisfied/loyal customers tend to join the loyalty scheme and reap the benefits that the scheme provides. Examining the membership moderation effect, we found a consistent pattern across our four studies that loyalty programme membership does not significantly influence any of the links between CE, satisfaction and loyalty. These results indicate that by adopting a loyalty programme, retailers can target the right customers, i.e. the more satisfied and loyal customers, and possibly ‘maintain’ their loyalty, but not ‘enhance’/’strengthen’ it. This is an
important finding, not only because the results are surprising to many researchers and marketing managers (Brown and Dant, 2008), but also because results that do not confirm expectations are crucial to scientific progress (Fanelli, 2012, Knight, 2003).

Third, we apply the technique of internal or ‘built-in’, multiple, differentiated replications, addressing the problem of the lack of replication research in literature (Ehrenberg, 1990, Evanschitzky et al., 2007, Hunter, 2001, Lindsay and Ehrenberg, 1993, Schmidt, 2009). We collected data in four separate and independent tranches from different cities and varied retail settings. This allowed us to arrive at the broad conclusions that we have reached, which would have been impossible with a single study based on a single set of data. Internal replication may also have the additional benefit of providing avenues with which to approach meta-analysis.

Managerial implications

This research has important practical implications. Managing experiences requires a thorough understanding of the critical attributes that set customer expectations, contribute to a customer’s experience, and ultimately connect to customer satisfaction and loyalty in various retailing contexts. Using this knowledge, managers can orchestrate a coherent and integrated series of marketing interventions that collectively meet customer needs and expectations. Managing the customer experience need not be overly complex because many of the expectations of customer service are highly functional, for example, merchandise quality ranks No. 1 in importance that forms retail CE in three out of our four studies. Particularly, for supermarket retailers in less-developed markets more emphasis should be put on product choice, price, and promotional activities which will make good customer experience as revealed from our data collected from a Tier 2 city (Xiamen). In addition, there are attributes such as service personnel, atmosphere, store exterior and interior design that
touch on the more human attributes that tend to address emotions rather than pure reason. Supermarkets and high-end department stores in the developed markets in particular need to focus on staging experiences of fun and fantasy, providing aesthetic, educational, entertaining and escapism values (Pine and Gilmore, 2011). Overall, retail companies must consider and manage the key attributes, from functional to emotional, to provide a holistic CE that involves ‘sense, feel, think, act and relate’ (Kumar et al., 2013).

While it is evident from this research that customer loyalty programmes do in fact ‘work’ in technical terms, e.g. they target satisfied and loyal market segments, enrol many members, and then reward the most valuable customers, the evidence nonetheless is that the incremental enhancements in customer loyalty may be very small. Managers therefore risk ‘giving the product away’ unnecessarily because customer loyalty remains virtually unaltered by membership in a loyalty programme. Today, loyalty programmes have become nearly ubiquitous among major service providers, which may be viewed as an industry-standard offering to customers that simply implies parity between competitors. This may explain their weak effect since they are regarded both by customers and providers as a standard part of service delivery and thus a cost of doing business (Shugan, 2005). In this respect, managers need to be realistic about the possible effects of loyalty programmes, which are often over-hyped (Keiningham, 2006).

Limitations and future research

This research needs to be interpreted considering the limitations of the research design. All four studies here employed cross-sectional surveys. Further research may use a combination of cross-sectional surveys with company databases to reduce such biases. The use of company databases containing behavioural data will also ameliorate the gaps that may exist between reported behaviour and actual behaviour. The sample sizes of many individual
competitors in the market surveyed were relatively small which prevented us from conducting meaningful comparisons at individual programme level. Further studies may examine whether our findings hold true for each individual programme, given the possible heterogeneous nature of the programmes on offer in the study, and for different segments of the market. Finally, we would like to echo Evanschitzky and Armstrong’s (2012) call for more studies based on multiple, in-built, differentiated replication studies, and Fanelli’s (2012) call for measures to encourage publication of negative results, particularly those contradict important predictions and/or previous positive evidence so as to ensure self-correction of the literature in the long run.
References


Fornell, C. and Larcker, D.F. (1981), "Evaluating structural equation models with


Ringle, C.M., Wende, S. and Will, A. (2005), "SmartPLS 2.0 (M3)", in, SmartPLS, Hamburg


Schmitt, B.H. (1999), "Experiential marketing: How to get customers to sense, feel, think, act and relate to your company and brand", *New York.*


Tynan, C. and McKechnie, S. (2009), "Experience Marketing: a review and reassessment",


Figure 1 Conceptual Model
### Table 1 Scale Items & Convergent Validity

<table>
<thead>
<tr>
<th>Constructs/Items</th>
<th>Study 1 (n=315)</th>
<th></th>
<th></th>
<th>Study 2 (n=211)</th>
<th></th>
<th></th>
<th>Study 3 (n=241)</th>
<th></th>
<th></th>
<th>Study 4 (n=300)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>VIF</td>
<td>Weight</td>
<td>Loading</td>
<td>VIF</td>
<td>Weight</td>
<td>Loading</td>
<td>VIF</td>
<td>Weight</td>
<td>Loading</td>
<td>VIF</td>
<td>Weight</td>
</tr>
<tr>
<td>Customer experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merchandise quality</td>
<td>1.2</td>
<td>.47**</td>
<td>.72**</td>
<td>1.41</td>
<td>.48**</td>
<td>.72**</td>
<td>2.29</td>
<td>0.13</td>
<td>.47**</td>
<td>1.56</td>
<td>.31**</td>
</tr>
<tr>
<td>Service personnel</td>
<td>1.32</td>
<td>.33**</td>
<td>.71**</td>
<td>1.93</td>
<td>-.07</td>
<td>.17</td>
<td>1.82</td>
<td>.30**</td>
<td>.80**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Store atmosphere</td>
<td>1.31</td>
<td>.32*</td>
<td>.68**</td>
<td>2.15</td>
<td>-.15</td>
<td>.05</td>
<td>2.88</td>
<td>.24*</td>
<td>.48**</td>
<td>2.14</td>
<td>0.08</td>
</tr>
<tr>
<td>Value for money</td>
<td>1.2</td>
<td>.25*</td>
<td>.61**</td>
<td>1.48</td>
<td>.37*</td>
<td>.70**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>1.11</td>
<td>.13</td>
<td>.41**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.73</td>
<td>.24*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td></td>
<td>.17</td>
<td>.23*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Promotional activities</td>
<td>1.86</td>
<td>.23*</td>
<td>.71**</td>
<td>1.27</td>
<td>0.05</td>
<td>.79**</td>
<td>1.83</td>
<td>0.14</td>
<td>.66**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choice of products</td>
<td>1.74</td>
<td>.29**</td>
<td>.46**</td>
<td>2.66</td>
<td>0.16</td>
<td>.77**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retailer reputation</td>
<td></td>
<td>1.5</td>
<td>.24**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Store facilities</td>
<td>2.29</td>
<td>0.16</td>
<td>.86**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choice of big brands</td>
<td>2.16</td>
<td>0.01</td>
<td>.66**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Store exterior</td>
<td></td>
<td>2.05</td>
<td>.22*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Store layout</td>
<td>2.12</td>
<td>.25*</td>
<td>.80**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product presentation</td>
<td></td>
<td>2.39</td>
<td>.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfaction</td>
<td></td>
<td>CR=.77</td>
<td>AVE=.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAT1</td>
<td>.87**</td>
<td>.82**</td>
<td>.83**</td>
<td>.85**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAT2</td>
<td>.70**</td>
<td>.87**</td>
<td>.88**</td>
<td>.88**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loyalty</td>
<td></td>
<td>CR=.85</td>
<td>AVE=.66</td>
<td>CR=.84</td>
<td>AVE=.73</td>
<td>CR=.84</td>
<td>AVE=.73</td>
<td>CR=.84</td>
<td>AVE=.73</td>
<td>CR=.86</td>
<td>AVE=.75</td>
</tr>
<tr>
<td>LOY1</td>
<td>.85**</td>
<td>.76**</td>
<td>.71**</td>
<td>.83**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOY2</td>
<td>.85**</td>
<td>.88**</td>
<td>.75**</td>
<td>.65**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOY3</td>
<td>.74**</td>
<td>.84**</td>
<td>.80**</td>
<td>.59**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: *p < .05; **p < .01; VIF = variance inflation factor, CR = composite reliability, AVE = Average variance extracted.
Table 2 Latent Variable Correlations & Square Roots of AVE

<table>
<thead>
<tr>
<th></th>
<th>CE</th>
<th>SAT</th>
<th>LOY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAT</td>
<td>.47</td>
<td>.70</td>
<td>.81</td>
</tr>
<tr>
<td>LOY</td>
<td>.53</td>
<td>.68</td>
<td></td>
</tr>
<tr>
<td>Study 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAT</td>
<td>.64</td>
<td>.84</td>
<td></td>
</tr>
<tr>
<td>LOY</td>
<td>.61</td>
<td>.60</td>
<td>.82</td>
</tr>
<tr>
<td>Study 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAT</td>
<td>.80</td>
<td>.85</td>
<td>.75</td>
</tr>
<tr>
<td>LOY</td>
<td>.47</td>
<td>.41</td>
<td></td>
</tr>
<tr>
<td>Study 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAT</td>
<td>.82</td>
<td>.87</td>
<td></td>
</tr>
<tr>
<td>LOY</td>
<td>.55</td>
<td>.59</td>
<td>.70</td>
</tr>
</tbody>
</table>

Notes: Boldface numbers on the diagonal are the square root of the average variance extracted. CE=customer experience, SAT=satisfaction, LOY=loyalty, NA= not applicable.
<table>
<thead>
<tr>
<th></th>
<th>Beijing (n=106)</th>
<th>Shanghai (n=107)</th>
<th>Shenzhen (n=102)</th>
<th>Total sample (n=315)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M1</td>
<td>M2</td>
<td>M1</td>
<td>M2</td>
</tr>
<tr>
<td>CE → SAT</td>
<td>.50**</td>
<td>.49**</td>
<td>.42**</td>
<td>.37**</td>
</tr>
<tr>
<td>SAT → LOY</td>
<td>.61**</td>
<td>.60**</td>
<td>.59**</td>
<td>.54**</td>
</tr>
<tr>
<td>MEM → SAT</td>
<td>.21**</td>
<td>.04</td>
<td>.21**</td>
<td>.03</td>
</tr>
<tr>
<td>MEM → LOY</td>
<td>.07</td>
<td>.02</td>
<td>.07</td>
<td>.08</td>
</tr>
<tr>
<td>CE x MEM → SAT</td>
<td>.22</td>
<td>-.18</td>
<td>.22</td>
<td>.14</td>
</tr>
<tr>
<td>SAT x MEM → LOY</td>
<td>-.08</td>
<td>-.08</td>
<td>-.08</td>
<td>-.20</td>
</tr>
<tr>
<td>R²(SAT)</td>
<td>.29</td>
<td>.34</td>
<td>.23</td>
<td>.31</td>
</tr>
<tr>
<td>∆ R²(SAT)</td>
<td>.05</td>
<td>.08</td>
<td>.05</td>
<td>.08</td>
</tr>
<tr>
<td>R²(LOY)</td>
<td>.39</td>
<td>.39</td>
<td>.40</td>
<td>.40</td>
</tr>
<tr>
<td>∆ R²(LOY)</td>
<td>.00</td>
<td>.00</td>
<td>.03</td>
<td>.03</td>
</tr>
</tbody>
</table>

Notes: *p<.05, **p<.01; M1 = Model 1 (main model), M2 = Model 2 (with moderators); CE = customer experience, SAT = satisfaction, LOY = loyalty, MEM = loyalty programme membership. Control variables: age, gender, education and income.
Table 4  Structural Model – Study 2-4

<table>
<thead>
<tr>
<th></th>
<th>Study 2 (n=211)</th>
<th>Study 3 (n=241)</th>
<th>Study 4 (n=300)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M1</td>
<td>M2</td>
<td>M1</td>
</tr>
<tr>
<td>CE → SAT</td>
<td>.63**</td>
<td>.65**</td>
<td>.80**</td>
</tr>
<tr>
<td>SAT → LOY</td>
<td>.57**</td>
<td>.50**</td>
<td>.44**</td>
</tr>
<tr>
<td>MEM → SAT</td>
<td>-.11*</td>
<td>-.07</td>
<td>.07</td>
</tr>
<tr>
<td>MEM → LOY</td>
<td>.23**</td>
<td>-.03</td>
<td>.03</td>
</tr>
<tr>
<td>CE x MEM → SAT</td>
<td>-.19</td>
<td>.11</td>
<td>-.03</td>
</tr>
<tr>
<td>SAT x MEM → LOY</td>
<td>-.09</td>
<td>-.16</td>
<td>-.01</td>
</tr>
<tr>
<td>( R^2(\text{SAT}) )</td>
<td>.43</td>
<td>.46</td>
<td>.66</td>
</tr>
<tr>
<td>( \Delta R^2(\text{SAT}) )</td>
<td>.03</td>
<td>.01</td>
<td>.01</td>
</tr>
<tr>
<td>( R^2(\text{LOY}) )</td>
<td>.42</td>
<td>.48</td>
<td>.22</td>
</tr>
<tr>
<td>( \Delta R^2(\text{LOY}) )</td>
<td>.06</td>
<td>.03</td>
<td>.00</td>
</tr>
</tbody>
</table>

Notes: *p<.05, **p<.01; M1 = Model 1(main model), M2=Model 2 (with moderators); CE=customer experience, SAT=satisfaction, LOY=loyalty, MEM=loyalty programme membership. Control variables: age, gender, education and income.