

Development and Inspection of WeChat Electronic Service Quality Scale

WANG Wei^{[a],*}; WU Jianping^[b]

^[a] Postgraduate, Department of Business Management, Jiangsu University of Science and Technology, Zhenjiang, China.

^[b] Associate professor, Jiangsu University of Science and Technology, Zhenjiang, China. Research direction is marketing.

*Corresponding author.

Received 23 August 2014; accepted 29 October 2014

Published online 30 November 2014

Abstract

In this paper, WeChat Electronic Service Quality is taken as the object of study. By the mainstream scale development paradigm, it adopted methods of questionnaire survey, factor analysis and etc, and the MIM-E-QUAL scale which includes 6 dimensions such as Reliability, Assurance, Personalization, Enjoyment, Perceived price level and Ease of use were developed; in addition, AMOS21.0 confirmatory factor analysis (CFA) was used for Scale Validation. As a result, the scale is reliable and effective.

Key words: Electronic Service Quality (E-SQ); WeChat; Scale development; Scale validation

Wang, W., & Wu, J. P. (2014). Development and Inspection of WeChat Electronic Service Quality Scale. *International Business and Management*, 9(2), 92-99. Available from: <http://www.cscanada.net/index.php/ibm/article/view/5879>
DOI: <http://dx.doi.org/10.3968/5879>

INTRODUCTION

WeChat is a kind of Mobile Instant Messenger (hereinafter MIM for short) based on the sociality relationship and used for communication network, mobile payment, information sharing, games, etc. At present, the total

users of WeChat has been exceeded 600 million in which the active users have reaches 480 million. So WeChat is taking up the dominant position of domestic MIM. But in the network times which is full of fierce competition, the product and service provided by each manufacturers are more and more homogeneous, substitute products emerge endlessly, and the sticky improvement of electronic users are more difficult than traditional services (van Riel et al., 2001). So improving service quality is the key point for the MIM suppliers to attract users and avoid invalid users and to win the competition advantage.

In recent years, some foreign scholars do some related research on the basic testing tools of service quality level—Electronic Service Quality Scale (Electronic Service Quality for short). But most of the researchers focus on internet shopping (or can be called on line retail, electronic retail), the research focus on E-SQ are not common, the E-SQ of the mobile client are seldom mentioned either. We can see that the research of the electronic service quality of mobile instant message is still in a fledging period. So the development of a kind of WeChat electronic service quality considering fitting for the specific culture background of our country makes important sense to evaluate electronic service quality and improve service quality level.

This article bases on this kind of MIM software, focus on users' perception, learns from the research result of traditional and electronic service, combines with the environment and characters of WeChat, uses E-SQ scale development widely used questionnaire inquiry method to collect data, follows the scale development paradigm which was advanced by Churchill in 1979 to develop WeChat electronic service quality scale: MIM-E-QUAL and verify it.

1. MEASUREMENT OF ELECTRONIC SERVICE QUALITY

1.1 Theoretical Basis for the Establishment of E-SQ Measurement Scale

1.1.1 E-SQ Measurement Scale Is the Reorganization, Revision and Extension of the Traditional Service Measurement Scale

Among the various researches on service quality, the SERVQUAL scale proposed by Parasuraman et al. in 1998 was the most representative traditional service quality scale with the most extensive influence containing five dimensions including tangibles, reliability, responsiveness, empathy and assurance. Some of the scholars believed that the key point for the measurement of e-service quality was to differentiate the on-line service from the traditional service. Therefore in order to be more adapted to the online trading environment, the SERVQUAL scale should be re-modified and re-organized with the reference of the background characteristics of the online trading (Santos, 2003). In view of this, when these scholars were measuring service quality, they focused mainly on the revision and reconstruction of the typical SERVQUAL scale. For example, O’Niell et al. (2001) had applied additionally Contact to the original SERVQUAL that had been designed with three dimensions, including Responsiveness, Reliability and Tangibles when the quality of the online book service was measured. While Gefen (2002) had combined Customer Trust, Perceived Risk with vendor, Perceived Risk with vendor and Customer loyalty with the original five dimensions in SERVQUAL to constitute a composite service quality for the purpose to study the relationship between the service quality of the e-commerce website and the customer loyalty.

1.1.2 It’s Necessary to Re-develop E-SQ Scale According to Its Characteristics and Attributes

The other scholars held that the scale should be re-organized according to the characteristics of the electronic service with the reference of different attributes under different environments. The representative scales that had been worked out included: SITEQUAL, WEBQUAL, e-TailQ, E-S-Qual, E-Res-QUAL, eTransQual and e-SELFQUAL etc. Took “e-TailQ” scale as an instance, it was developed by Wolfinbarger and Gilly (2003) based on the concept of Total Quality Management with the application of three survey methods through the online and offline focused interview team, literature processing team and online customer team to measure and study the customer-perceived online retail service quality. There were totally 14 items contained in this scale to forecast the customer’s e-SAT, their loyalty and attitude. Meanwhile the scale had been designed with four main dimensions, including Web site design, reliability/fulfillment, privacy/security and customer service.

1.2 E-SQ Measurement Method

Many of the scholars have mainly adopted qualitative research method, quantitative research method and the combination of qualitative and quantitative research method in the E-SQ measurement (Ladhari, 2010). The qualitative research method had been adopted mainly in the earlier researches. For example, with the adoption of the literature research method, Zeithaml (2000) had summarized and developed an E-SQ scale covered with 11 dimensions including Reliability, Ease of navigation, Price knowledge, Access, Efficiency, Personalization, Security/privacy, Site aesthetics, Responsiveness, Assurance/trust and Flexibility on the basis of numerous documents. In fact, similar methods had been applied by Madu and Madu (2002) as well.

Quantitative research method, which has been adopted by most of the researchers during their studies, has been widely applied in the E-SQ measurement. Through the investigation on the 416 users who had purchased or used the web services in the recent 3 months, Cristobal et al. (2007) had constructed a scale with four dimensions: web design, customer service, assurance and order management. Also in 2005, through a random sampling of some users who had the experience in online shopping, Parasuraman and Zeithaml managed to develop an E-S-QUAL scale containing two parts, say regular service contact and service failure (E-ResS-QUAL) by adding some technology-related dimensions to the original SERVQUAL scale. In the first part, it was designed with 4 core dimensions, including efficiency, system availability, fulfillment and privacy. While in the second part, which was E-ResS-QUAL scale, there were three dimensions, including compensation, responsiveness and contact.

However there’re also some scholars, who would like to develop scales with the combination of both qualitative and quantitative research methods. For example, with the reference of the comments made by the consumers on their online experiences as brokers, Yang and Fang (2004), who had integrated the content analysis on the key events related to online banking services with the online services survey to analyze two online consumer review websites (Gomez.com and Epinions.com), had collected 848 consumers’ information and data to obtain 17 service quality dimensions according to their analysis. All of these dimensions had been classified into three categories, including customer service quality, online system quality and product or service variety.

The opinions proposed by Yang and Jun (2002) and Lai (2006) et al., are agreed with in this paper, holding that it’s not applicable to copy the traditional services during the measurement of e-service quality. Scales will be reorganized with reference of the traditional service after the combination of the environment and the characteristics of the electronic service.

2. THE ESTABLISHMENT OF THE MIM-E-QUAL SCALE

2.1 Generation of Initial Scale

Nowadays, the development paradigm of the measuring tools that has been widely recognized and accepted in the field of management was proposed by Churchill (1979), who suggested that scales should be developed according to the following four steps: (1) Definition of concepts; (2) design of the dimensions and the initial items in the scale; (3) data collection; (4) refinement of the measuring items. According to this theoretical basis, first sort out and make a conclusion on the numerous predecessors' documents. Second, organize a small-scale seminar, where there're

eight persons present mainly consisting of students and teachers who are using WeChat frequently and were asked to fill out the first draft of the questionnaire on the spot for the purpose to make an evaluation on the following questions about the testing range, the service content, the understandability of the questions and the veracity of the narration. Finally according to the feedback information from the seminar members and on account of the problems found during their filling out the questionnaire, a reasonable adjustment was made on the first draft of the questionnaire. In this way, the initial 9 dimensions and 35 measuring items were obtained in this paper on the basis of the said work with the specific definition shown in Table 1.

Table1
Definitions of the Dimension

Dimension	Definition
Reliability	Ability of the WeChat operators to provide accurate and stable services
Assurance	User confidence in the WeChat operators itself and the overall strength of the provided services
Privacy/Security	The safety of user's personal information, property information, and privacy
Personalization	The ability WeChat provided to meet the design, function, experience and other individual needs for different users
Communication	Users take part in communications and social activities by WeChat
Enjoyment	Emotional perception triggered by the user's experience of electronic services that WeChat provided
Perceived price level	User perception about internet connection fees and the cost of virtual services cost by WeChat
Interface design	The visual and the comfort experience about interface design
Ease of use	How can the users easily and quickly use the app

2.2 Pretest

Before the formal questionnaire was issued, items were processed for the preparation of the pre-test questionnaire. Then through the internet, totally 45 questionnaires were issued to some of my classmates and friends who had been using WeChat frequently with 31 copies of the valid questionnaires having been received to realize an effective recovery rate of 68%. After that, the scale was refined through the test of CITC and the Cronbach's Alpha for the pretest questionnaire through SPSS19.0. After the removal of Item RE1 and RE2, whose CITC values were separately $0.478 < 0.5$ and $0.433 < 0.5$, the Cronbach's Alpha of the whole dimension was thus increased to be 0.790 from 0.774. Due to the improvement of α coefficient, both of the items were then deleted. Just in the same way, it's easy to find that there's a significant improvement in α coefficient after Item AS4 and Item EN3 have been deleted. It turns out that the scale would be featured with high reliability if the Cronbach's Alpha of all dimensions is > 0.7 after the adjustment.

2.3 Issuance of Questionnaire and Data Collection

Questionnaires have been issued through the combination of internet and mobile phone in a large sample survey. When both of the methods are adopted, the questionnaire

is usually prepared on the website of SOJUMP to work out the online questionnaire, which is then linked to the cell phone by scanning the QR code through WeChat after the design of the questionnaire. Then a mobile phone-based questionnaire is generated, which would then be issued through WeChat. According to the survey by iResearch, it showed that about 80% of the WeChat users were those people below 35 years old, especially those between 25 to 30 years old (Nov. 2013). In view of this, the students and those white-collar employees who haven't been working for many years and still prepare to change their job any time are chosen as the objects in this paper. By virtue of the central database of a recruitment website in Beijing, 100 job applicants were chosen randomly for the web-based questionnaire, which was then issued to them by email. As to the cell phone-based questionnaire, it was started from the author's classmates and friends to send the questionnaire directly to the circle of friends through WeChat. Then the questionnaire would be forwarded by means of snowball sampling. Totally 100 copies of the web-based questionnaires were issued with 67 copies having been received. While for the cell-phone based questionnaire, totally 121 copies were received with the valid questionnaires up to 164 copies finally after the rejection of 24 invalid questionnaires.

2.4 Scale Purification

After data collection, first made a reliability analysis on the pretest questionnaire, where the scale was refined through SPSS19.0, which was applied as well through the pretest method for the test of the CITC and Cronbach's Alpha. The method won't be repeated herein. Two items, PS3 and PE3 were deleted. Second, KMO and Bartlett's Test proved that the 29 items were not independent from each other in essence. There was a correlation between them, which made it suitable for the application of factor analysis. With the adoption of the principal component analysis covered in the factor analysis, those factors whose characteristic root was greater than 1 were extracted. Then 0.5 was set as the intercept point to choose a load factor. Therefore those items where any of the load factor was below 0.5 or the multiple load factors were greater than 0.5, were deleted. Since the load factors on the first and the sixth factor in RE5 were separately 0.47 and 0.467, both of which were below 0.5 and failed to conform to the requirements, then RE5 was deleted. In addition to this, the following items were removed as well for the same reason: RE5, CO1, CO2 and EO3.

Finally, factor analysis was conducted again on the rest 25 items with two of such iterations having been made additionally. Each time, those items whose load factor was lower than 0.5, were deleted, when it's required that the multiple load factors should be greater than 0.5. Then in the last factor analysis, 6 factors for 22 items were extracted with each dimension compliant with the requirement. It turned out that the cumulative variance contribution was 73.809%, which had guaranteed a favorable result.

Table 2
Load Factor of Itmes

Dimension	Item	Component					
		1	2	3	4	5	6
F1	RE4	-	-	-	-	-	0.813
	AS1	0.857	-	-	-	-	-
	AS2	0.850	-	-	-	-	-
F2	AS3	0.843	-	-	-	-	-
	PS1	0.726	-	-	-	-	-
	PS2	0.651	-	-	-	-	-
	PS4	0.831	-	-	-	-	-

To be continued

Continued

Dimension	Item	Component					
		1	2	3	4	5	6
F3	PE1	-	0.776	-	-	-	-
	PE2	-	0.759	-	-	-	-
	PE4	-	0.755	-	-	-	-
F1	CO3	-	-	-	-	-	0.759
	EN1	-	-	-	-	0.744	-
F4	EN2	-	-	0.304	-	0.614	-
	EN4	-	-	-	-	0.674	-
	PR1	-	-	0.747	-	-	-
F5	PR2	-	-	0.775	-	-	-
	PR3	-	-	0.759	-	-	-
	PR4	0.355	-	0.667	-	-	-
	ID2	-	-	-	0.735	-	-
F6	ID3	-	-	-	0.744	-	-
	EO1	-	-	-	0.849	-	-
	EO2	-	-	-	0.799	-	-

Note. Extraction Method: Principal component analysis
 Rotation Method: Varimax with Kaiser Normalization
 a. Rotation converged in 8 iterations

2.5 Ultimate Dimension

F1 contained two items: RE4 and CO3, both represented the ability to provide services accurately, so named this dimension: reliability. F2 contained six items: AS1, AS2, AS3, PS1, PS2, PS4. This six were involved in the trust to WeChat operators and service, including personal information, so named this dimension: assurance. F6 contained four items: ID1, ID3, EO1, EO2, all represented the convenience of interface design or function design, so named this dimension: Ease of use. The rest kept the original dimension name.

Table 3
The Item of MIM-E-QUAL Scale

Dimension	Number	Initial number	Item	Name
F1	K1	RE4	I can communicate and entertain through WeChat any time and any where	Reliability
	K2	CO3	I am sure to succeed in logining WeChat	
	B1	AS1	I trust in the operator of WeChat	
	B2	AS2	I trust in WeChat	
F2	B3	AS3	I trust in the services of WeChat	Assurance
	B4	PS1	My personal and property information is safe	
	B5	PS2	My Chat records are safe	
	B6	PS4	Only if I set it myself, or I will not be found by strangers or get messages from them	
F3	G1	PE1	Logining WeChat in my own mobile phone can make me very comfortable	Personalization
	G2	PE2	The system will automatically recognize my account	
	G3	PE4	According to preference, I can change chat background and interface	
F4	Y1	EN1	I'm glad to chat with friends through WeChat	Enjoyment
	Y2	EN2	I'm glad to browse information through WeChat	
	Y3	EN4	WeChat make me pleased	
F5	Z1	PR1	I only need to pay for data charge and do not need to pay for additional cost of communication	Perceived price level
	Z2	PR2	I can afford data charge	
	Z3	PR3	Content price (Sticker gallery, game prop etc.) of WeChat is appropriate	
	Z4	PR4	I will buy this paid content if needed	
F6	J1	ID2	Navigation and interface of WeChat make me feel convenient	Ease of use
	J2	ID3	The interface design match with my mobile phone screen	
	J3	EO1	Functions of WeChat (Shake, drift bottle, scan QR code etc.) are easy to operate	
	J4	EO2	With just a few clicks, I can find informations and services I want	

3. VERIFICATION OF SCALE

Although simplified items for the scale have been obtained, it's still necessary to test the validity and stability of the scale. On account of this, confirmatory factor analysis has been applied in this paper to verify further the scale through AMOS21.0.

3.1 Content Validity

Content validity, which is also called as face validity, refers to the logical conformity between the measurement concept and content. Generally subjective judgment is applicable to content validity. In this article, the dimension chosen has been centered on the predecessors' research achievements with multiple verifications having been made. Meanwhile items have been selected with the reference of the scales described in the classical documents. In fact, after the completion of the initial

dimensions, the selection of items have been evaluated repeatedly through the organization of the seminars and pretests to put the missing items back and delete those improper items for the purpose to adjust the logical structure to guarantee the rationality of the items in the questionnaire. Therefore this scale is featured with reasonable content validity.

3.2 Convergent Validity and Discriminant Validity

Convergent validity has been used to explain the high correlation between a theoretical concept measured by people and the different measuring tools in the same structure, showing that such measuring tools are characterized with convergent validity. In this article, the skew is between - 0.757 and 0.253, whose absolute value is below 1, while the kurtosis is between - 0.737 and 1.737, whose absolute value is below 2. All of these reflect that the data conforms to normal distribution.

Therefore it's feasible to apply Maximum likelihood to estimate parameters with CMNI = 259.361, DF = 192, CMNI/DF = 1.351 < 3 and RMSEA = 0.033 < 0.06. In fact, except NFI = 0.824, which is below the standard criticalvalue of 0.9, all of the rest values, IFI, TLI and CFI are separately 0.947, 0.935 and 0.946, which are all more than 0.9. Since there's little difference between NFI and the critical value, the result is still acceptable on account of the great influence exerted by the sample

size. From the above, it's easy to come to a conclusion that the model has been characterized with high imitation degree. Actually as shown in the figure, the standardized regression weights of these 22 observed variables are between 0.61 and 0.82, which are all above the critical value of 0.6, indicating that there's a strong correlation between each observed variable and the corresponding latent variable, proving that the index has been featured with high convergent validity.

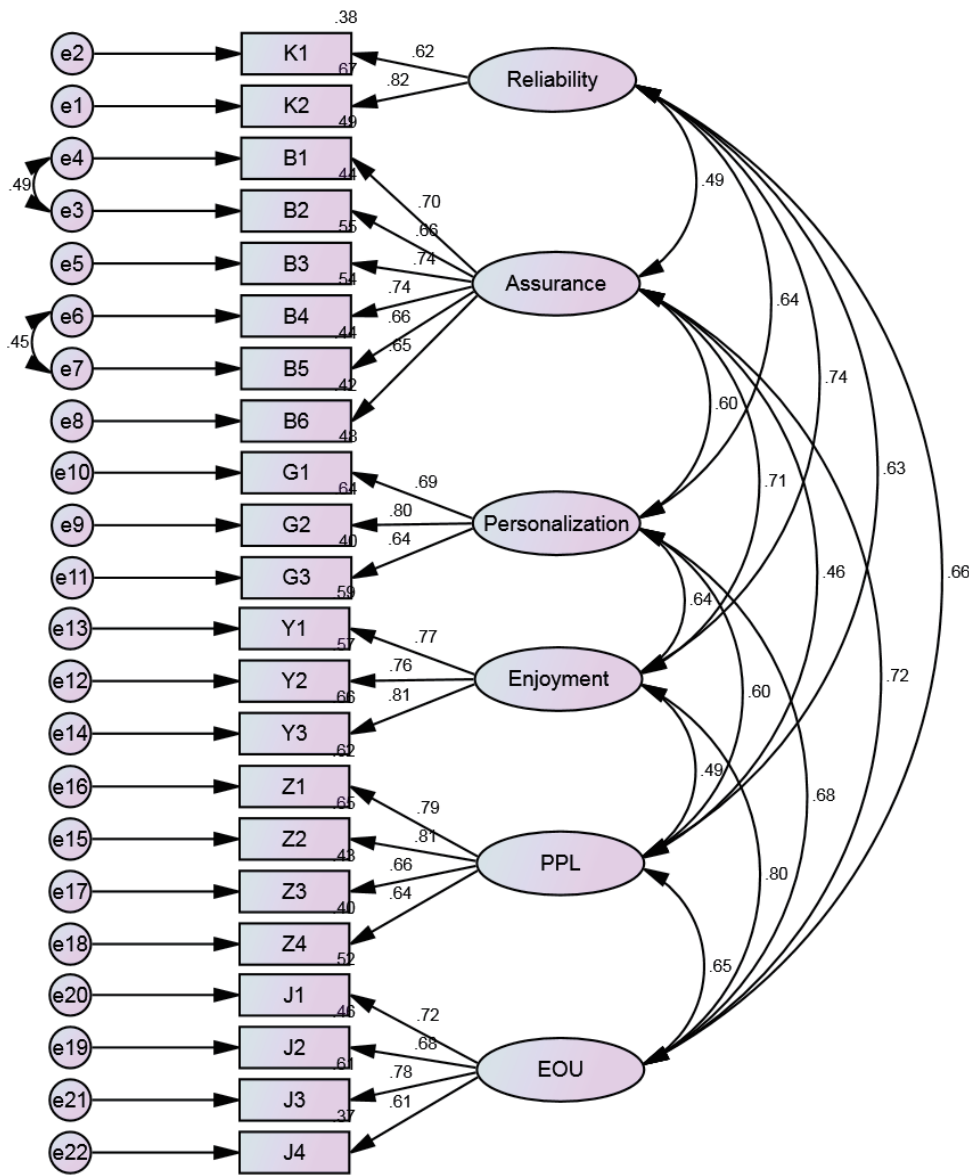


Figure1
First-order CFA Model

Note. PPL: Perceived price level; EOU: Ease of use

Discriminant validity has been used to explain the low correlation between the scale and the measuring tools in different concepts, showing that such scale is featured with discriminant validity. The discriminant validity of a dimension can be observed through the factor score

weights. As shown in the table, the score weight of these condary index for a dimension (the part marked in bold in Table 4) is apparently higher than that of the other dimensions, showing that the scale is featured with ideal discriminant validity.

Table 4
Factor Score Weights

Item/Dimension	Ease of use	Perceived price level	Enjoyment	Personalization	Assurance	Reliability
J1	0.134	0.019	0.370	0.014	0.032	0.013
J2	0.119	0.016	0.330	0.012	0.029	0.011
J3	0.178	0.025	0.490	0.019	0.043	0.017
J4	0.080	0.011	0.220	0.008	0.019	0.008
Z1	0.030	0.181	- 0.008	0.020	0.005	0.042
Z2	0.036	0.215	- 0.010	0.024	0.006	0.050
Z3	0.016	0.098	- 0.005	0.011	0.003	0.023
Z4	0.017	0.100	- 0.005	0.011	0.003	0.023
Y1	0.039	- 0.006	0.193	0.014	0.040	0.062
Y2	0.036	- 0.005	0.174	0.012	0.036	0.055
Y3	0.054	- 0.007	0.249	0.018	0.051	0.079
G1	0.014	0.012	0.013	0.144	0.017	0.024
G2	0.026	0.023	0.024	0.267	0.032	0.045
G3	0.012	0.011	0.011	0.123	0.015	0.021
B1	0.016	0.001	0.018	0.008	0.142	- 0.005
B2	0.011	0.001	0.013	0.006	0.099	- 0.004
B3	0.024	0.002	0.027	0.013	0.214	- 0.008
B4	0.017	0.002	0.019	0.009	0.149	- 0.006
B5	0.009	0.001	0.010	0.005	0.079	- 0.003
B6	0.012	0.001	0.014	0.006	0.109	- 0.004
K1	0.007	0.015	0.033	0.014	- 0.006	0.167
K2	0.018	0.036	0.079	0.034	- 0.015	0.405

3.3 The Second-order CFA Model

In order to identify the effectiveness of a dimension, the second-order CFA model, which is an exception of the first-order CFA model, has been applied in this article. Generally it's believed that the value of a factor loading is between 0.50 and 0.95. According to the output of AMOS, it's easy to find that the standardized estimates of the various dimensions in Table 3 and MIM-SQ are between 0.676 and 0.910 with C. R. value above 1.96 and *P* less than 0.001, showing that the model has been featured with excellent goodness of fit with high convergent validity.

Table 5
Standardized Regression Weights

Path	Estimate	S. E.	C. R.	P
Reliability←MIM-SQ	0.788	0.184	5.794	***
Assurance←MIM-SQ	0.734	0.170	5.401	***
Personalization←MIM-SQ	0.777	0.162	5.918	***
Enjoyment←MIM-SQ	0.871	0.183	6.205	***
Perceived pricelevel←MIM-SQ	0.676	0.161	5.334	***
Ease of use←MIM-SQ	0.910			

Note. ***represents $P < 0.001$.

For Ease of use, the path coefficient is assumed to be 1. Therefore, there's no output content under the last three items.

Meanwhile it's applicable to test the discriminant validity of a dimension through the MTMM matrix (Multitrait-Multimethod). As shown in Table 6, it's apparent that the reliability (the part marked in bold

in table 6)of each dimension is higher than the other correlation coefficients of the dimensions, reflecting that all of these six dimensions have been featured with excellent discriminant validity.

Table 6
Mimm Matrix

Dimension	Ease of use	Perceived price level	Enjoyment	Personalization	Assurance	Reliability
Ease of Use	0.828					
Perceived Price Level	0.615	0.540				
Enjoyment	0.793	0.493	0.759			
Personalization	0.708	0.526	0.677	0.604		
Assurance	0.668	0.496	0.639	0.578	0.539	
Reliability	0.717	0.533	0.686	0.411	0.411	0.621

CONCLUSION

Based on the users' perception perspective, this article provides a WeChat E-SQ frame, including 6 dimensions and 22 items: Reliability, Assurance, Personalization, Enjoyment, Perceived price level, Ease of use. Thereinto, reliability stands for the service provider ability to provide service, contains 2 items; assurance stands for the trust of users, contains 6 items; personalization stands for users' individual needs, contains 3 items; enjoyment stands for users sense appreciation, contains 3 items; perceived price level stands for users' value appreciation, contains 4 items; ease of use stands for users' operation convenience, contains 4 items. Using empirical approach method to refine and inspect the scale, verifies that it can measure E-SQ reliably and validity, and can provide reference for future E-SQ research. In the future research, this scale can be used in other MIM applications or expand to other field of mobile electronic service, to verify the scale repeatedly and continually so that to improve the validity and scientificity. It can also be used to do the research on the service result of perceived price level, satisfaction and loyalty. It is the trend and direction of E-SQ research.

REFERENCES

Churchill, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, 2(16), 64-73.

Cristobal, E., Flavian, C., & Guinaliu, M. (2007). Perceived e-service quality (PeSQ): Measurement validation and effects on consumer satisfaction and web site loyalty. *Managing Service Quality*, 17(3), 317-340.

Gefen, D. (2002). Customer loyalty in e-commerce. *Journal of the Association for Information Systems*, 3, 27-51.

Ladhari, R. (2011). Developing e-service quality scales: A literature review. *Journal of Retailing and Consumer Services*, 17, 464-477.

O'Neill, M., Wright, C., & Fitz, F. (2001). Quality evaluation in on-line service environments: An application of the importance-performance measurement technique. *Managing Service Quality*, 11(6), 402-417.

Parasuraman, A., Zeithaml, V. A., & Malhotra A. (2005). E-S-Qual: A multiple-item scale for assessing electronic service quality. *Journal of Service Research*, 7(3), 213-233.

Parasuraman, A., Zeithaml, V. A., & Berry, L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of services quality. *Journal of Retailing*, 64(1), 12-40.

Santos, J. (2003). E-service quality: A model of virtual service quality demensions. *Management Service Quality*, 13(3), 233-246.

Van Riel, A. C. R., Liljander, V., & Jurriëns, P. (2001). Exploring consumer evaluations of e-services: A portal site. *International Journal of Service Industry Management*, 12(4), 359-77.

Wolfenbarger, M. & Gilly, M. C. (2003). ETailQ: Dimensionalizing, measuring and predicting retail quality. *Journal of Retailing*, 79(3), 183-198.

Yang, Z., & Fang, X. (2004). Online service quality dimensions and their relationships with satisfaction: A content analysis of customer reviews of securities brokerage services. *International Journal of Service Industry Management*, 15(3), 302-326.

Zeithaml, V. A., Parasuraman, A., & Malhotra, A. (2000). E-service quality: Definition, dimensions and conceptual model. *Working Paper*, Marketing Science Institute, Cambridge, MA.