Empirical Research on a Model to Measure End-User Satisfaction with the Quality of Database Search Results

Ma Feicheng, Mo Zuying and Luo Yi

(Center Studies of Information Resources, Wuhan University)

Abstract: We constructed an end-user model to measure end-user satisfaction with the quality of database search results, using the customer satisfaction theory as a metric. We investigated end-user satisfaction and analyzed key factors which affected user satisfaction. The results show that the end-users' perception of value is the key factor among all of the different factors that impact satisfaction with regard to quality and end-users are willing to make efforts to obtain a higher quality of data. Users tend to evaluate their satisfaction from the perspective of their demands, and database developers should be user oriented in order to improve the level of satisfaction with the data in the database.

Keywords: Database quality; end-user satisfaction model; quality of search results;

1. Introduction

An academic, periodical database is the cornerstone of innovation in the scientific research community. The quality of information provided by a database directly affects innovation. Therefore, it is important for database developers to improve the level of end-user satisfaction in order to maintain a competitive advantage and capture a larger share of the market.

There is a lot of research on customer satisfaction at present (Ahmad A. et al., 2012; Fang Y.H. et al., 2011; Joseph C. et al., 2000; Yuanquan Li et al., 2007), especially in the fields of physical products and services. Meanwhile, customer satisfaction models and assessment instruments have been implemented, which are widely used in many fields. Research of customer satisfaction, mainly focuses on commercial and service industries, which consists of three aspects. The first one is concerned with traditional service industries (Assaf A.G., & Magnini V., 2012;

Lecic-Cvetkovic D. et al., 2012), such as hotel management, postal services, stock markets, commercial banks, etc. They evaluate the level of customer satisfaction and analyze the various factors which probably impact customer satisfaction by empirical studies. The second aspect deals with the research on the quality of the websites (Chia-Lin Hsu et al., 2012; McKinney V. et al., 2002; Lin C.C. et al., 2011; Dan J. Kim, 2012; Hanne Sorum, 2011), such as online shopping websites, booking websites, tourism information service center websites, etc. The research aims to study the relationships between the quality of the website and customer satisfaction. The third aspect concentrates on research of service quality (Ooi K.B. et al., 2011; Lee W.I., & Lee C.L., 2011; Kyriakopoulos G.L., 2011; Yuanquan Li et al., 2007; Lee H.S., 2010), studying how the quality of the service affects customer satisfaction and evaluating the level of customer satisfaction in regard to online services, mobile communication services, etc.

In recent years, scholars in the field of library and information sciences in China have begun to apply the customer satisfaction model in order to evaluate databases. Liren Gan et al. (2004) constructed an ACSI-based, evaluation framework for website databases, and measured the end-user satisfaction level with four, famous website databases; Li Li et al. (2007) developed an integrated framework for measuring satisfaction based on the characteristics of the end users, and discovered disconfirmation as a key construct among other factors influencing end-user satisfaction; Qiong Tang and Xinhe Zhang (2007) proposed an end-user satisfaction based model which was applied to test end-user satisfaction with the databases of Sun Yat-Sen University Library and gave recommendations on how to develop digital collections of Sun Yat-Sen University Library. Shunli Lei (2010) discussed the relationship between end-user satisfaction and their behavior based on the analysis of the variations between the different levels of satisfaction with users who use library at different frequencies, and found that end-user satisfaction is positively related to their behavior. These studies mainly make an overall evaluation of databases, websites or library collections by using the satisfaction model, and the evaluated aspects cover the quality of database, service, system and so on. However, there is a lack of research on

the end-user satisfaction model of retrieving results from the database. The quality of database search results in the database is user oriented, and it directly influences end-user satisfaction with the quality of information in the database. The purpose of this research is to investigate the level of end-user satisfaction with the quality of the search results, and analyze key factors impacting the level of satisfaction.

2. Theoretical Background

2.1 The model of customer satisfaction

Early research about customer satisfaction mainly focuses on the disconfirmation model, studying the relationship between customer expectation, perception, and satisfaction. Oliver (Oliver R.L., 1980; Oliver R.L., 1999) proposed the *Expectancy-Disconfirmation Theory*. The term expectation refers to what consumers believe they should and will receive from sellers through any given transaction. After a period of initial consumption, customers form perceptions about a sellers' performance, and assess their perceived performance, vis-à-vis, their original expectation and determine the extent to which their expectation is confirmed. They form a satisfaction level based on their disconfirmation and their expectations. Westbrook (Westbrook R.A., & Reilly M.D., 1983; Westbrook R.A., 1980), on the other hand, proposed an alternative theoretical approach, the *Value-Percept Disparity Model*, which, through empirical examinations, directly opposes the expectation disconfirmation theory. They found that there is a need for continued efforts to improve the measurement of theoretical constructs in order to determine the level of consumer satisfaction. As the research delved deeper, "Disconfirmation" is considered the result produced by comparing with a certain standard. Oliver (Oliver R.L., 2010) summarized the comparing standards and corresponding psychological percepts, such as the expectation, customer need, the ideal performance, equity, and so on.

Nowadays, the disconfirmation of expectations paradigm is most widely used in evaluating the level of customer satisfaction, and has become the main evaluation

model of the SCSB (Swedish Customer Satisfaction Barometer), the ACSI (American Customer Satisfaction Index), and the ECSI (European Customer Satisfaction Index), etc. The ACSI (Fornell C. et al., 1996) represents a new type of customer-based measurement system for evaluating and enhancing the performance of firms, industries, economic sectors, and national economies. It measures the quality of goods and services as experienced by the customers that consume them. It represents a cumulative evaluation of a firm's market offering, rather than a person's evaluation of a specific transaction. Factors involved in the ACSI include customer expectations, the perceived level of quality and value, the customer's overall satisfaction, customer complaints and customer loyalty. In this study, we will develop the ACSI model, and apply it to evaluate database end-user satisfaction with the quality of retrieved information.

2.2 Categories and dimensions of the quality of information

There is a great deal of research on the various categories and dimensions of the quality of information (Ballou D.P., & Pazer H.L., 1985; Laudon K.C., 1986; Richard Y. Wang et al., 1995; Thomas C. Redman, 1996; Richard Y. Wang, 1998; Beverly K. Kahn et al., 2002;). Ballou and Pazer divided data quality into four dimensions based on various attributes of information: accuracy, timeliness, consistency, and completeness. Richard considers data as an information product, and defined the quality of information as a fitness for use by information consumers. He divided data quality into four categories and fifteen dimensions: intrinsic IQ, accessibility IQ, contextual IQ, and representational IQ. Beverly et al. constructed the PSP/IQ (Product and Service Performance/ Information Quality) model, which assigned two aspects of quality, *conforming to specifications*, and *meeting or exceeding consumer expectations*, as columns of PSP/IQ model; for two rows of PSP/IQ model, they chose *product quality* and *service quality*, and thus formed four quadrants in the PSP/IQ model. The four types of data are sound information, dependable information, useful information, and usable information.

In China, the quality and dimensions of digital databases are studied by many scholars (Li Li et al, 2009; Xianjin Zha, & Minghong Chen, 2010; Youhua Liu et al., 2008; Jingjuan Zhao et al., 2009; Yuling Li et al., 2009). Li Li et al. studied the quality of data on websites, and created several metrics to measure database quality: reliability, document variety, novelty, timeliness, and information system quality, that includes intelligibility, functionality, accessibility, etc. Xianjin Zha et al. proposed a set of practical assessment indicators, which categorizes the quality of information into the quality of content, the quality of form, the quality of the system, and the quality of the usefulness of the database. With various views offered by different researchers, it's hard to form a general standard for the quality of information that can suit all research fields. However, it is necessary and feasible to construct an objective quality system of database within a certain field. This study takes academic, periodical, full-text database as the subject. In order to obtain the users' view of the quality of the data in database, we randomly selected 10 graduate students who were using the database frequently and conducted face-to-face interviews. The results of the interviews includes the users' understanding of information quality, key quality points that were germane users' research in the database, and key factors affecting the quality of data in the database. We used open questions in the interviews. By interviewing users, we find what concerns users most about the quality of data is whether they can obtain the data they need and the quality of the content. It also shows that the quality of the content in the database mainly involves the completeness, the authoritativeness, and the novelty of the content, from a users' perspective. Therefore, we studied the quality of the data in the database using four dimensions, namely, completeness, authoritativeness, novelty, and the quality of the search results. According to the interviews, we set some evaluation indicators for each dimension, as shown in Table 1. Completeness refers to the completion degree of database collection and the completeness of resources, including resources quantity, time span, coverage of a subject and the number of complete texts. Authoritativeness refers to the depth and credibility of the literature, including the ratio of core journals, the average frequency of cited articles, and the authoritativeness of an article's content. Novelty

refers to the time lag that exists between the inclusion of articles in the database and the novelty of an article's content as perceived by users. Quality of search results refers to whether the user can retrieve the needed resources, and it includes two major aspects.First is the retrieval performance, with commonly used indices being: recall ratio and precision ratio. Second is the usability of the descriptive information provided by each retrieval record, such as whether the descriptive information can help users to effectively judge the relativity or content quality of the literature.

Among the quality indicators in Table 1, some are objective, while others are subjective, such as “recall ratio”, or “precision ratio”, which can be computed by using a formula. In this study, we will investigate how users perceive and score each indicator, so the quality indicators in this paper are subjective as perceived by users.

Table 1. database quality dimensions and indicators

Quality	Quality dimensions	Quality indicators
Quality of Data	Completeness	Resources quantity; Time span; Full-text ratio; Coverage of a subject;
	Authoritativeness	Ratio of core journal; Average frequency of cited articles; Authoritativeness of an article's content;
	Novelty	Novelty of an article content; Time-lag collected by database;
	Quality ofsearch results	Recall ratio; precision ratio; Descriptive information used to judge relativity; Descriptive information about quality;

In this paper, we mainly study end-user satisfaction with the quality of data from search results. We constructed the satisfaction model based on the investigation of the users of the CNKI, periodical, full-text, database (the CNKI, full-text database is one of the most popular academic periodical databases in China.), and analyzed the key factors that impact end-users' satisfaction with the quality of data.

3. The Conceptual Model

The models of customer satisfaction are mostly based on the Disconfirmation

Paradigm, which is used to evaluate the level of customer satisfaction by the disconfirmation between the expected quality and the perceived quality. In this section, we will construct a conceptual model which illustrates end-user satisfaction with the quality of a database's search results, based on the ACSI model and the dimensions and indicators of the quality of database displayed in Table 1. We adopted the "Expectancy-Disconfirmation Theory", and compared quality of information users expectations to the quality they perceived, evaluating the level of end-user satisfaction. We also considered the cost of obtaining data, and introduced the concept of the perception of value, which refers to the ratio of the quality of data the user obtained and the costs incurred. According to the quality and indicators of search results, end-user satisfaction includes both the satisfaction with the overall quality and satisfaction with the quality of the search results. The purpose of this paper is to study how the end-users' satisfaction with the quality was impacted, but we did not factor in customer complaints and loyalty as considered in the ACSI model. The conceptual model of database end-user satisfaction with the quality of the data retrieved as search results mainly include:

(1) End-users' perception of quality End-users' perception of quality refers to end-users' perception and judgment of the quality of the database search results in the process of using the database; it is an important factor impacting end-user satisfaction. The perceived quality, as an endogenous latent variable, cannot be measured directly, so we use the quality indicators of search results as observation variables to evaluate it. The observation variables of the perceived quality are end-users' perception of the recall ratio, precision ratio, descriptive information used to judge relativity, and descriptive information about quality.

(2) End-users' expectation of quality as the standard for comparison End-users' expectation of quality refers to an ideal state of the quality of data that users hope to achieve according to their current demand and past experience. It is comprised of two aspects: the expectation of how much the quality can meet individual demand, and the other is the expectation of quality generated by past experience. The expected quality, as the latent variable, cannot be measured directly, so we use the quality

indicators of search results as observation variables to evaluate it. The expected quality is described by the importance of quality indicators users perceived, so the observation variables of expected quality consist of the weight of recall ratio, precision ratio, descriptive information used to judge relativity, and descriptive information about quality that users perceived.

(3) End-user perception of value The end-users' perception of value refers to the ratio of quality data that users obtained and the cost users incurred. In this research, our respondents are university users who can freely obtain data from the database. The costs users incurred are mainly the efforts they make to search the information, including their time and energy. So the perceived value in the model is the ratio of the quality of data retrieved from the search results and the efforts made to search the results. It had two observation variables: the efforts made to search the information relative to the quality of data retrieved from the search results (the perceived value 1), and the quality of data retrieved from the search results relative to the fore-mentioned efforts (the perceived value 2).

(4) End-user satisfaction End-user satisfaction refers to end-users satisfaction or dissatisfaction after using the database to obtain data. In this paper, it specifically refers to users' comprehensive satisfaction formed according to the performance of the database during the last three months. User satisfaction includes the overall satisfaction with the quality and the satisfaction with the quality of the search results. The satisfaction with the search results, as observation variable, refers to end-user quality satisfaction with the search results of the database, while the overall quality satisfaction as latent variable refers to the satisfaction with the overall quality of database in database that users perceived. The observation variables of the overall quality satisfaction cover the quality satisfaction relative to end-user demand, the quality satisfaction relative to end-user expectation, and the overall evaluation of quality.

We choose observation variables and latent variables of the Structural Equation Model to describe the causal relationship between the elements of the conceptual model. The conceptual model of end-user satisfaction with the quality of the data

retrieved is shown in Figure 1. The variables in rectangle are observation variables. The variables in ellipse are latent variables. The direct lines between latent variables and observation variables means that latent variables are represented by the observation variables; The direct lines between latent variables means that they have a causal relationship resulting from users perceptions and their preferences, and the independent variables point to dependent variables. In the conceptual model, expected quality is an exogenous latent variable, and the perceived quality, perceived value, the level of satisfaction with search results, and the overall satisfaction are endogenous variables; the exogenous latent variables impact endogenous variables, and endogenous variables have correlations with each other. The relationships between latent variables (including "quality satisfaction with search results") are represented by the directed thick lines, and we suppose all of them are positive correlation.

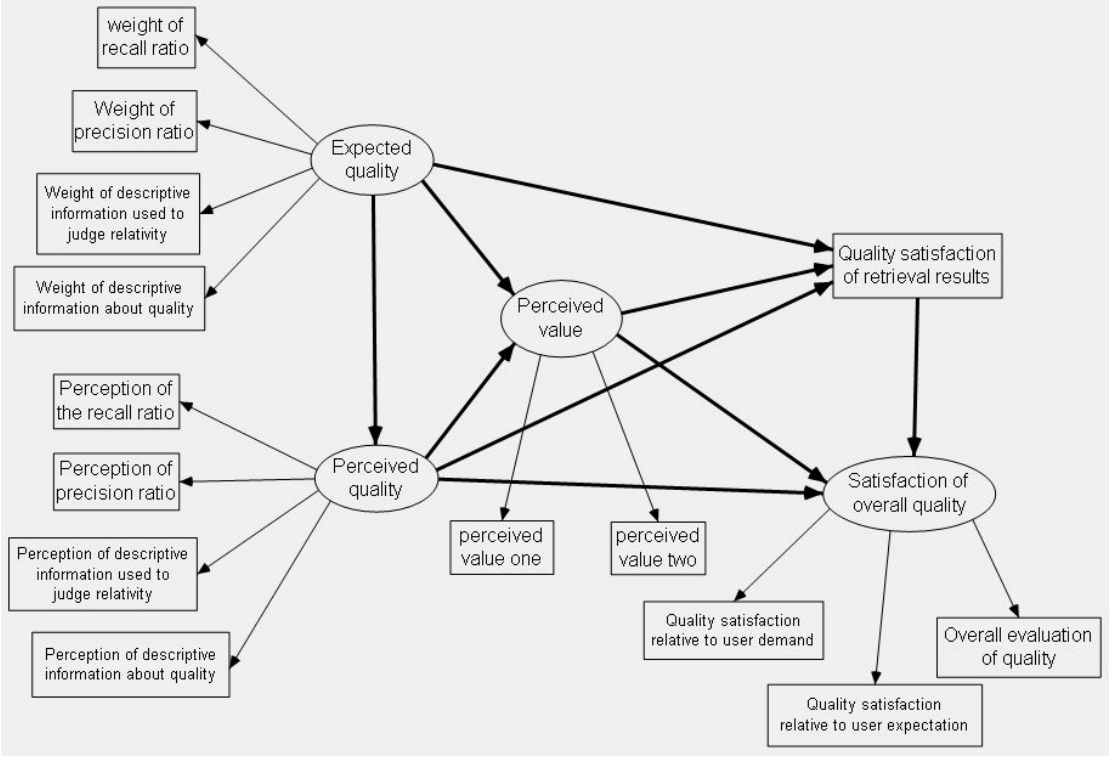

Figure 1. Conceptual model of end-user satisfaction with the quality of search results

The list of latent variables and observation variables in the conceptual model is shown in Table 2. Adopting Likert Level 5 Scales, we surveyed database users by questionnaire in order to find end-user perception of observation variables. The items in the questionnaire were designed according to observation variables, and each

observation variable is designed as an item in the questionnaire.

Table 2 Structure variables and Observational variables in conceptual model

Structure variables	Observational variables	
Expected quality	Expected quality of search results	Weight of recall ratio; Weight of precision ratio; Weight of descriptive information used to judge relativity; Weight of descriptive information about quality;
Perceived quality	Perceived quality of search results	Perception of the recall ratio; Perception of precision ratio; Perception of descriptive information used to judge relativity; Perception of descriptive information about quality;
Perceived value	Efforts relative to the quality of database search results that user accessed (the perceived value 1);	
	the quality of database search results that user accessed relative to the efforts (the perceived value 2);	
Satisfaction with quality dimension	Quality satisfaction with search results;	
Satisfaction with overall quality	Quality satisfaction relative to user demand;	
	Quality satisfaction relative to user expectation;	
	Overall evaluation of quality;	

4 Research methodology and data collection

This paper takes China National Knowledge Internet (CNKI) academic journal full-text database as an example. CNKI is one of the world's largest, full-text databases of Chinese academic journals, which is continuously updated. The content covers science, engineering, agriculture, philosophy, medicine, the humanities and

social sciences etc. The CNKI included 7686 kinds of domestic academic journals as of October 2010; the core journal rate is 96%. It included fixed number of year since 1915, and part of journal date back to the start publication. We collect data by questionnaires, and use the structural equation model (SEM) and software to do parameter estimation and model verification. Meanwhile it analyzes the influence factors and mechanism of the end-user satisfaction for the quality of database search results, and thus constructs a end-user quality satisfaction model.

4.1 Study samples and data collection

First of all, we completed the design of the questionnaire based on the review of literature and group discussion. The questionnaire included three parts and 14 items, which makes for a one-to-one correlation with the observation variables in the concept model. The first part is about the investigation on end-users' expectations of quality; the second part is about the investigation on end-user quality cognitive; the third part is about the investigation on end-user value perception, their satisfaction for search results quality and the overall satisfaction. The questionnaire is set in Likert 5 Scale. The link is as follows: "http://www.sojump.com/jq/1624491.aspx".

Secondly, we sampled the users of our survey. The respondents in this investigation were researchers in the university, including postgraduate students, doctoral-postgraduate students, and postdoctoral students. They were the main users of the CNKI journal full-text database, using it frequently, freely and stably via the university library. The respondents were chosen randomly.

Thirdly, in order to collect data, we distributed questionnaires randomly via the internet and conducted face-to-face interviews with questionnaires. The 'face-to-face interviews with questionnaires' was the distribution of paper questionnaires to the participant cohort, which was different from distributing a digital questionnaire via the internet. Among the respondents were students from universities in Hubei, Henan, Beijing, Guangdong and other 13 provinces in China. Altogether 320 questionnaires were distributed and 286 questionnaires were returned, among which 268

questionnaires were valid, the response rate was 89.375%; the proportion that was valid was 93.706%.

4.2 The Structural Equation Model

In terms of the research methodology, we used the structural equation model to construct end-users' satisfaction model. The structural equation model is a research method based on statistical analysis technology, which can be used to deal with the complex multi-variable research data. Adopting such a model, we can achieve the following functions (Haidong Wu, 2009): handling multiple interdependent variables, allowing the dependent and independent variables containing measurement errors, estimating factor relationships and the fitting degree between the model and the sample data, and dealing with more flexible and complicated measurement models. It is very suitable for research of social sciences, whose indicies could not be measured accurately.

In this paper, we used AMOS 17.0 software, which is specially suited for the Structural Equation Model (SEM), to analyze and verify the model. AMOS uses a kind of covariance structure analysis, combined with the traditional, general, linear model and common factor analysis technology. It can be used to do all kinds of SEM model analysis. The structural equation model (Minglong Wu, 2009) can handle the overall model fit and pays attention to the comparison of the overall model. In the overall model fit test, we need to test the differences between two matrices, the overall covariance matrix (matrix) and the covariance matrix (· · matrix) of hidden variables in the hypothetical model. The null hypothesis is: matrix = . . matrix. Because in reality, we couldn't get either the variance or the covariance of the overall data, or parameters from overall data, we use covariance matrix and the parameters derived from sample data instead of overall data. In the model fit test, we use the significant level of 0.05, namely if significance test probability value $p<0.05$, the significant level will be reached, and the null hypothesis will be refused, that is, there is significant difference between covariance

matrix of sample data and covariance matrix of hidden variables in the hypothesis model, and the two are not well-fitted. Conversely, if the significance test probability value $p > 0.05$, the difference will not be significant and the null hypothesis will be accepted, which means the hypothesis model and the sample data fit one another well. We use the maximum likelihood estimation method for the parameter estimation according to the data sample size, normality and independence.

5 Data Analysis and the Results

5.1 The profile of the respondents

Of the 268 valid questionnaires, 64.2% of the respondents were postgraduate students, and 35.8% were doctoral-postgraduate students and those of a higher degree. The ratio of postgraduate students is obviously higher than that of the doctoral-postgraduate students and those of a higher degree. The main reason for this deviation is that the number of doctoral-postgraduate students and students of a higher degree is significantly smaller than that of postgraduate students in our country. In terms of the respondents' subject distribution, the students of liberal arts account for 55.6%, slightly higher than the percentage of the students of science and engineering, which amounts to44.4%. As we may see, the two percentages are balanced basically. With regard to the respondents' experiences using the CNKI database, 70.5% of respondents have used the database for more than three years, and 69.7% of the respondents use the database more than three times per week, which means most of them have used the CNKI database many times and they are familiar with it. As to the respondents' search skills, they have received professional training regarding database information retrieval during their undergraduate study, so we assume that the respondents' search skills in our investigation are consistent and they can express their cognition for dimensions of quality accurately.

5.2 The reliability and validity tests

In the field of social sciences, the Cronbach coefficient (α), which is also called internal consistency coefficient, is used to estimate the reliability of Likert scale. According to the 4 latent variables in the conceptual model corresponding to 14 observational variables, we divide them into 4 groups and inspect the overall reliability of each group. The observational variable 'the satisfaction for search results quality is grouped with three observational variables as 'end-user satisfaction with quality'. The Cronbach coefficients of 4 groups are shown in Table 3 and they are all greater than the standard of 0.7. The result shows that the observational variables for each group have a high overall reliability and a good internal consistency.

In the reliability test, if the reliability coefficient becomes much higher when a variable is deleted, this variable may be different from other variables and it can be deleted in data analysis. The 'Alpha if Item Deleted' of the 14 variables is shown in Table 3, the data value shows that all 'Alpha if Item Deleted' is smaller than the Cronbach coefficient, thus all variables pass the reliability test.

Table 3. Reliability Test

Group (latent variables)	Cronbach coefficient (α)	observational variables	Alpha(α) if Item Deleted
Expected quality	0.839	Weight of recall ratio	0.817
		Weight of precision ratio	0.813
		Weight of descriptive information used to judge relativity	0.776
		Weight of descriptive information about quality	0.778
Perceived quality	0.702	Perception of the recall ratio	0.672
		Perception of precision ratio	0.598
		Perception of descriptive information	0.600

		used to judge relativity	
		Perception of descriptive information about quality	0.683
Perceived value	0.702	the perceived value 1	---
		the perceived value 2	---
end-users' satisfaction with quality	0.857	Quality satisfaction with search results	0.844
		Quality satisfaction relative to user demand	0.804
		Quality satisfaction relative to user expectation	0.809
		Overall evaluation of quality	0.814

In the validity test (Minglong Wu, 2010), validity can be divided into content validity, criterion-related validity, construct validity and construct validity can test the extent of theoretical construction of the psychological traits. The questionnaire used in this paper is constructed on end-user satisfaction model and the theory of the information resource quality, so it is suitable to use the construct validity to test validity. Since we can work out the construct validity with the factor analysis, we used it to do the validity test. Before the factor analysis is made, the calculation for variables' KMO test values is needed to ensure that the sample data is suitable for the analysis. We used SPSS 19.0 to calculate the variables' KMO test value, which is 0.899, greater than the standard of 0.7 given by statistician Kaiser (Liren Gan et al., 2010), and Bartlett spherical test value reached the significant level of 0.01 which shows the sample data is suitable for the factor analysis.

The component matrix in Table 4 shows the factor loading of each variable, and represents the degree of correlation of variables and common factors. Seen from the Table, all variables pass the validity test except the 'Perception of descriptive information used to judge relativity'; its factor loading (0.43) is less than the standard value of 0.45 and is considered to be deleted. The other variables pass the validity test, showing that the behavior or psychological trait they measured has consistency.

Table 4. Component Matrix

Variables	Component 1	Variables	Component 1
the perceived value 2	.764	Perception of descriptive information used to judge relativity	.643
Quality satisfaction relative to user demand	.747	Perception of the recall ratio	.629
Overall evaluation of quality	.745	Weight of precision ratio	.546
the perceived value 1	.733	Weight of descriptive information about quality	.539
Quality satisfaction relative to user expectation	.720	Weight of descriptive information used to judge relativity	.526
Quality satisfaction with search results	.704	Weight of recall ratio	.465
Perception of precision ratio	.647	Perception of descriptive information about quality	.430

According to the reliability and validity test results, we abandoned the variable of "Perception of descriptive information used to judge relativity", and deleted it in the conceptual model (Figure 1) before the model test.

5.3 The structural equation model

We ran AMOS 17.0 to construct the conceptual model map of satisfaction for the quality of results of information resource search and then imported data to do model path analysis and calculation. After adjusting the conceptual model many times, we ultimately got a model of satisfaction for the quality of results of information resource search which fits the sample data well.

The standardized path coefficient diagram of the search results quality satisfaction model is shown in Figure 2. According to the model fit indices on the right top of the diagram, the chi-square value is 67.335, the significant probability p is 0.089>0.05, AGFI (adjusted goodness-of-fit index), GFI (goodness-of-fit index) are greater than 0.9, RMSEA (the root mean square error of approximation) =0.032<0.05. These indicators fit that model very well. The path coefficient between latent variables and observation variables in the figure is the factor loadings, which shows the relative

importance of the observation variable to the latent variable. Seen from the path coefficient values, the factor loadings are between 0.68 and 0.89, indicating a well-fitted model and observational variables can effectively reflect the trait of measured constructs. The bold Italic digitals in the diagram are the variables' multiple correlation square values which represent the explained variance. It is usually regarded as the estimated minimum value of reliability; the value is greater than 0.5, which shows the inner quality inspection of that model is good. In the model, only the latent variable 'perceived quality' is less than 0.5, which may result from the incompleteness of its observation variables. Other variables are either close to or more than 0.5, which indicates that the model works well.

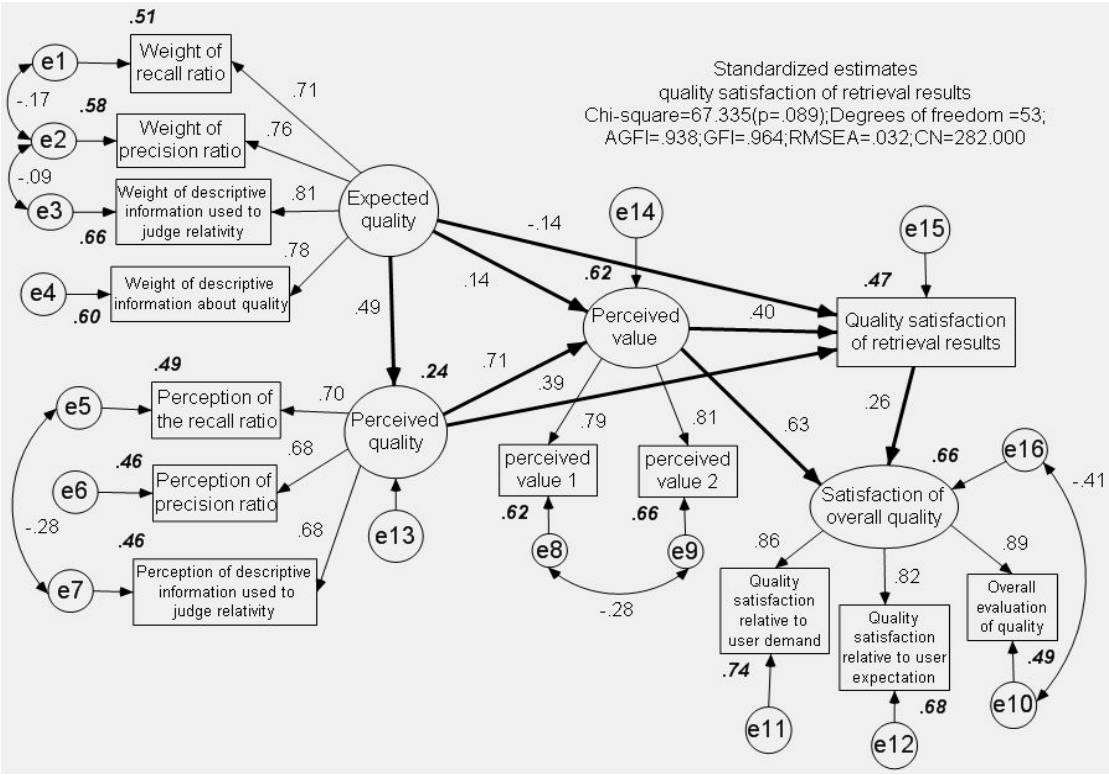

Figure 2. Standardized path coefficient diagram of satisfaction model of search results quality

In Figure 2, the number on path between latent variables (including 'the search results quality satisfaction') indicates the standardized regression coefficient of two variables; the numerical symbol indicates the positive or negative correlation between variables, and the number indicates the degree of influence. This path coefficient value needs undergoing the significant test to ensure the support from sample data. The test of the estimated value of the path coefficient (Minglong Wu, 2009) is

conducted to ensure whether the regression path coefficient estimation value is equal to 0. Table 5 displays the test results of non-standardized regression coefficient estimation value. In Table 5, critical ratio (C.R.) equals to the ratio of parameter estimate and standard error (S.E.), which is equivalent to the test value. If the absolute value of this ratio is greater than 1.96, the parameter estimation value reaches 0.05, and the absolute value of critical ratio is greater than 2.58, then the parameter estimation value will reach 0.01, which is on a significant level. Only paths whose parameter estimation value reaches the significant level could be supported by sample data, while others could not. We used p =0.05 as the significant level; '***' indicates the probability value of the significant level is less than 0.001, which means the sample data supports this path. According to the inspection data in Table 5, all paths between latent variables pass the test and gain support from sample data.

Table 5. Test results of regression coefficient estimation value

			Estimate	S.E.	C.R.	P
Perceived quality	<---	Expected quality	.413	.070	5.920	***
Perceived value	<---	Expected quality	.138	.070	1.961	.050
Perceived value	<---	Perceived quality	.835	.110	7.619	***
Quality satisfaction of search results	<---	Perceived value	.494	.170	2.908	.004
Quality satisfaction of search results	<---	Expected quality	-.170	.078	-2.174	.030
Quality	<	Perceived	.567	.19	2.98	.003

			Estimate	S.E.	C.R.	P
satisfaction of search results	<---	quality		0	4	
satisfaction with overall quality	<---	Perceived value	.697	.112	6.243	***
satisfaction with overall quality	<---	Quality satisfaction of search results	.230	.065	3.522	***

In the model fit test, due to the influence of the size of sample data on the chi-square value, other fit indices should also be considered. The major fit indices in this model are shown in Table 6, including absolute fit test index, such as CMIN/DF, P value, which is based on the covariance matrix hidden in hypothesis model and sample variance matrix (Haidong Wu, 2011), and the relative fit test index, such as GFI (goodness-of-fit index), RMSEA (the root mean square error of approximation), CFI (comparative fit index), and other indicators. Seen from data in the table, these indexes all reach the fit standard, which suggests the model fits the sample data well. Thus, both the regression coefficient of test results and data analysis have credibility.

Table 6. The major fit indices of the model

Fit Indices	Standards	Test results	Fit or not
CMIN/DF	<3	1.270	Yes
GFI	>0.9	.964	Yes
AGFI	>0.9	.938	Yes
RMSEA	<0.05	.032	Yes
PGFI	>0.5	.561	Yes
CFI	>0.9	.991	Yes
CAIC	Value of Conceptual model < Value of Saturated model; Value of Conceptual model < Value of Independence model;	317.793<599.780; 317.793<1713.426;	Yes

6. Discussion and Conclusions

With the end-user satisfaction model of the quality of the data retrieved from search results (Figure 2), we discussed and analyzed the relationship between the variables and the path coefficient, in order to find the key factors that impact end-users' satisfaction with the quality of the data retrieved.

(1) The analysis of path coefficient between latent variables and observed variables

According to the path coefficient value between the "expected quality" and its corresponding observed variables in Figure 2, we can learn that "the descriptive information used to judge relativity" is the most important variable of the latent variables. While, in the observed variables of "perceived quality", there is little distinction among the three indexes, which demonstrates the inconsistency between the expected quality and the real, perceived quality. It reflects that there exists an inconsistency between supply and demand. In the observed variables of the "perceived value", the relative importance of "perceived value 2" is more than that of "perceived value 1" (0.81>0.79), which means that users are willing to make efforts to gain data of a higher quality. In the observed variables of "overall satisfaction", the relative importance of "quality satisfaction relative to user expectation" is lower, which means that users tend to evaluate satisfaction from the perspective of their demands.

(2) The analysis of the relationship and the path coefficient value between the latent variables

We can infer from Figure 2 that the "overall satisfaction" is directly influenced by the "perceived value" and the "satisfaction with the quality of search results", while the influence of the former is more than that of the latter (0.63>0.26), and together they explain the variability of 66% of the overall satisfaction. The "satisfaction with the search results quality" gets little negative effect from the expected quality (-0.14), but receives positive effect from the perceived value and quality; the influence of the perceived value and perceived quality is nearly (0.40 vs.

0.39), which means that they are equally important to the satisfaction with the quality of search results, and together they explain the variability of 47%. The latent variable of “perceived value” is influenced by the expected quality and the perceived quality, and the degree of influence of the perceived quality is much more than the expected quality (0.71 vs. 0.14). The expected quality has an obvious influence on the perceived quality (0.49), so they both impact the perceived value and explain the variability of 62%.

The analysis above will practically promote the development of databases. Specifically speaking, first, database developers should be end-user oriented and provide search results needed by the users to improve end-users' satisfaction with the quality of the search results. Second, the perceived value is the key factor among all of the different influential factors regarding end-user satisfaction, and database developers should improve the end-users' perceived value of the search results of data by reducing the costs incurred and improving the quality of search results.

The quality of the data in the database can directly influence the ability of researchers to innovate, and determines the genesis of their innovations. This study has investigated the end-user satisfaction with the quality of information retrieved from the CNKI database, and constructed a model to measure the satisfaction of the end-user with the quality of database search results. It provides guidance for database developers to control the quality of data received via search results.

Acknowledgements

This research has been made possible through financial support by the Ministry of Education, Humanities and Social Science Foundation of China under Grant Number 12JJD820008 and the Natural Science Foundation of China under Grant Number 71173249. And we would like to give special thanks to Dongyan Zhang (Institute of Foreign languages, Wuhan University) for valuable assistance with the English language and composition.

References:

Ahmad A., Dey L., & Halawani S.M. (2012). A rule-based method for identifying the factor structure in customer satisfaction. *Information Sciences*, *198*, 118–129.

Assaf A.G., & Magnini V. (2012). Accounting for customer satisfaction in measuring hotel efficiency: Evidence from the US hotel industry. *International Journal of Hospitality Management*, *31*, 642–647.

Ballou D.P., & Pazer H.L. (1985). Modeling data and process quality in multi-input, multi-output information systems. *Management Science*, *31,* 150-162.

Beverly K. Kahn, Diane M. Strong, & Richard Y. Wang (2002). Information quality benchmarks: product and service performance. *Communications of the ACM, 45,* 184-192.

Chia-Lin Hsu, Kuo-Chien Chang, & Mu-Chen Chen (2012). The impact of website quality on customer satisfaction and purchase intention: perceived playfulness and perceived flow as mediators. *Inf Syst E-Bus Manage*, *10*, 549–570.

Dan J. Kim (2012). An investigation of the effect of online consumer trust on expectation, satisfaction, and post-expectation. *Information System and E-Business Manage*, *10*, 219-240.

Fang Y.H., Chiu C.M., & Wang E.T.G. (2011). Understanding customer satisfaction and repurchase intentions An integration of IS success model, trust, and justice. *Internet Research, 21*, 479–503.

Fornell C., Johnson M.D., Anderson E.W., Cha J.S., & Bryant B.E. (1996). The American Customer Satisfaction Index: Nature, Purpose, and Findings. *Journal of Marketing*, *60*, 7-18.

Haidong Wu (2009). An empirical study on the library end-user satisfaction based on structure equation modeling. *University of ChongQing in China,* ChongQing.

Haidong Wu (2011). The construction of evaluation indicator system of library reader satisfaction with structure equation modeling. *Information Science, 29,* 227-230.

Hanne Sorum (2011). An empirical investigation of user involvement, website

quality and perceived user satisfaction in e-government environments. *Electronic Government and the Information Systems Perspective, 6866,* 122-134

Jingjuan Zhao, & Huaiguo Zheng, et al. (2009). Evaluation studies on Chinese full-text database. *Journal of Modern Information, 29,* 162-165.

Joseph C., Michael K. Brady, & Tomas M. Hult (2000). Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments. *Journal of Retailing, 76,* 193-218.

Kyriakopoulos G.L. (2011). The role of quality management for effective implementation of customer satisfaction, customer consultation and self-assessment, within service quality schemes: A review. *African Journal of Business Management*, *5*, 4901–4915.

Laudon K.C. (1986). Data quality and due process in large inter-organizational records systems. *Communication of ACM, 29*, 4-11.

Lecic-Cvetkovic D., Kostic-Stankovic M., & Maslic V. (2012). Research of customer satisfaction with the electronic postal services in Serbia. *Metallurgy International*, *17*, 176–183.

Lee H.S. (2010). Effects of perceived value and service quality on customer satisfaction in the mobile phone service market. *Information-An International Interdisciplinary Journal*, *13*, 1207–1218.

Lee W.I., & Lee C.L. (2011). An innovative information and relationship between service quality, customer value, customer satisfaction, and purchase intention. *International Journal of Innovative Computing Information and Control*, *7*, 3571–3581.

Li Li, et al. (2007). The empirical study on satisfaction assessment of information-user for academic database websites. *Intelligence Journal, 26,* 546-554.

Li Li, Liren Gan, & Zhaoxia Xie (2009). Building satisfaction model for information users based on perceived quality of academic database websites. *Journal of the China Society for Scientific and Technical Information, 28,* 565-581.

Lin C.C., Wu H.Y., & Chang Y.F. (2011). The critical factors impact on online customer satisfaction. *World Conference on Information Technology (WCIT-2010), 3*,

276-281.

Liren Gan, et al. (2004). Evaluation of the user satisfaction level on four famous database websites. *Intelligence Journal, 23,* 524-530.

Liren Gan, et al. (2010). Construction and application of user satisfaction index model for library websites. *Journal of the China society for scientific and technical information, 29(2),* 159-168.

McKinney V., Yoon K., & Zahedi F. (2002). The Measurement of Web-Customer Satisfaction: An Expectation and Disconfirmation Approach. *Information systems research, 13*, 296–315.

Minglong Wu (2009). Structural equation model: Operation and Application of AMOS. *Published by ChongQing University Press,* ChongQing.

Minglong Wu (2010). Statistical analysis of questionnaire: Operation and Application of SPSS. *Published by ChongQing University Press,* ChongQing.

Oliver R.L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research, 17,* 460-469.

Oliver R.L. (1999). Expectation processes in satisfaction formation. *Journal of Service Research, 1,* 196-214.

Oliver R.L. (2010). Satisfaction: A behavioral perspective on the consumer. *Published by M.E. Shape, Inc.* New York 10504.

Ooi K.B., Lin B.S., Tan B.I., & Chong A.Y.L. (2011). Are TQM practices supporting customer satisfaction and service quality?. *Journal of Services Marketing*, *25*, 410–419.

Qiong Tang, & Xinhe Zhang (2007). A study of the quality evaluation model of library electronic resources based on user satisfaction. *Library Journal*, *26* ,46-53.

Richard Y. Wang, et al. (1995). Toward quality data: An attribute-based approach. *Decision Support Systems, 13,* 349-372.

Richard Y. Wang (1998). A product perspective on total data quality management. *Communications of the ACM, 41,* 58-65.

Shunli Lei (2010). Research on the application of model of university library resources evaluation based on user satisfaction. *Document, Information & Knowledge,*

136, 33-38.

Thomas C. Redman (1996). Data Quality for the Information Age. *Published by Artech House,* Boston, 1996

Westbrook R.A.(1980). Intra-personal affective influences on consumer satisfaction with products. *Journal of Customer Research*, *7*, 49-54.

Westbrook R.A., & Reilly M.D. (1983). Value-percept disparity: an alternative to the disconfirmation of expectations theory of consumer satisfaction. *Advances in consumer research*, *10*, 256-261.

Xianjin Zha, & Minghong Chen (2010). Research on quality assessment of information resources. *Journal of Library Science in China, 36,* 46-55.

Youhua Liu, & Aihua Qi, et al. (2008). The establishment and application of evaluation criteria systems for academic websites. *Information Science, 26,* 64-68.

Yuanquan Li, Jiayin Qi, & Huaying Shu (2007). A Study on Customer Satisfaction in Mobile Telecommunication Market by Using SEM and System Dynamic Method. *International Federation for Information Processing, 255*, 1221–1226.

Yuling Li, Xiangjun Chen, & Hui Li (2009). Fuzzy comprehensive evaluation for Chinese full-text periodical databases. *Information Science, 27,* 1844-1857.