Winter 12-10-2002

A Measurement Model of Trust in Internet Stores

Liping Liu

Cathy C. Li

Steven J. Karau

A Measurement Model of Trust in Internet Stores

Liping Liu
University of Akron
Department of Management
Akron, OH 44325-4801
liu@acm.org

Cathy C. Li
University of Akron
Daverio School of Accountancy
Akron, OH 44325-4802
cl16@uakron.edu

Steven J. Karau
Southern Illinois University
Department of Management
Carbondale, IL 62901-4627
skarau@cba.siu.edu

Abstract

In this study we theoretically develop and empirically test a measurement model of consumer trust in Internet stores. In particular, we define the notion of trust based on what is commonly agreed on by scholars across disciplines. We treat trust as a second-order construct and measure it using four first-order components: perceived security, privacy, integrity, and transactional accuracy. We conducted controlled experiments using three Internet bookstores and 173 subjects. We applied confirmatory factor analysis to determine the measurement efficacies.

1. Introduction

There have been many discussions on what propels successful Internet commerce (IC). Based on value-focused thinking, Keeney [31] suggested that IC success might be more a function of customer's belief and perception of the net value of the benefits and costs of both a product and the processes of finding, ordering, and receiving it. The Federal Administration and the Better Business Bureau showed that consumer's trust is a critical factor in stimulating Internet purchase. Castelfranchi and Tan [8] showed that a lack of trust is one of the main reasons that consumers and companies do not engage in IC. Keen [30] argued that the most significant long-term barrier to Internet purchasing will be the lack of consumer trust, both in the company's honesty and in the company's competence to fill Internet orders. A recent survey by Information Technology Association of America also found that 62% of respondents believed that trust was the top overall obstruct.

Since trust is an important aspect of IC success, the understanding of its meaning and measurement is imperative. For academic research, this construct can be used as a dependent or independent variable in a nomological network that links IC with preceding and ensuring constructs. For management practice, metrics are a way of learning what works and what does not, what is reinforcing and what is disconfirming feedback [49]. Unfortunately, there is an alarming lack of effort in validating the instrument for trust. A few studies have examined the notion of trust [9, 21, 27, 33]. However, their definitions and instruments of trust are different. It makes it difficult to compare and accumulate findings and thereby to develop syntheses of what is known [14]. Javenpaa et al. [27] felt a need to reexamine their instrument of trust adapted from marketing channels [17]. Lee and Turban [33] suggested that the construct should be reinvestigated in light of emerging technology, research and practice. They indicated that, if the construct is modeled incompletely and then integrated as part of a network of other constructs, the key aspects of variation among measures might be lost. The loss in explained variation can lead to errors in interpretation between dependent and independent variables within a system of path models [51]. The insufficiency of the metrics for IC trust corroborates with that of the metrics for information systems research in general and studies on net-enabled organizations in specific [53]. As Zmud and Boynton [58] noted, researchers have paid too little attention to measurement development issues and theoretical advancement has been constrained by the absence of reliable measures.

In this study we theoretically develop and empirically test a measurement model of consumer trust in Internet stores. In the parlance of latent variable statistics, we treat trust as a second-order construct and measure it using four first-order components: perceived security, perceived privacy, perceived integrity, and perceived transactional accuracy. The theoretical implication of higher-order models is that each first-order factor and the implied second-order factor are important in capturing the domain of the construct. Moreover, the second-order factor may be a more important mediator between a consequent and predictor variable than the first-order construct [51]. There is an additional consideration when choosing the current approach to the measurement of trust. Although it is prudent to borrow relevant scales from marketing to study IC, the unique feature of Internet technology warrants unique metrics for IC [52] and there is a danger associated with indiscriminate adoption [2]. In specific to the construct of trust in an Internet store, the unique feature of the Internet technology in terms of the conditions for trust to arise and the composition of trust [46] dictates that the notion of trust may be regarded as the synonym for security, privacy, or integrity; better cryptographic algorithms for data transmission and better authentication protocols for authenticity are considered equivalent to a higher level of trust. This consideration implies that trust in IC may be better reflected in multiple dimensions such as security, privacy, accuracy, and integrity.

2. A Theoretical Domain of Trust

To date, there is no universally accepted definition about what trust is. Economics define trust mainly as a phenomenon within and between institutions, and as the trust individuals put in those institutions [3]. Social psychology characterizes trust in terms of expectations and willingness of the trusting party in a transaction, the risks associated with acting on such expectations, and the contextual factors that either enhance or inhibit the development and maintenance of that trust [40]. Finance views trust as a level of subjective probability at which an agent will perform certain action [20].

Despite the differences in how trust is defined, scholars across disciplines seem to have two fundamental agreements. First, there is an agreement on what constitute trust: positive expectations of others [34] and willingness to be vulnerable [40]. Trust is a psychological state comprising the intention to accept vulnerability based on positive expectations of the behavior of another [46]. Second, why is someone willing to be vulnerable? There is an agreement on the necessary conditions for trust to arise: risk and interdependence. Risk is the perceived probability of loss [10, 35] and is an essential condition in psychological, sociological, and economic conceptualization of trust [16, 45, 57]. Trust would not be needed if actions could be undertaken without uncertainty and risk [35]. Uncertainty about what others intend to and will act appropriately is the source of risk [46]. Interdependence means the interests of one party cannot be achieved without reliance on another [46]. If one can achieve the interests without involving others, trust would not exist.

Table 1. Sample Trust Dimensions

Studies	Trust Dimensions
[4]	Information Security
[6]	Availability, Competence, Consistency, Discreetness, Fairness, Integrity, Loyalty, Openness, Promise Fulfillment, Receptivity
[15]	Ability, Trustworthy Intentions
[18]	Ability, Intention to delivery
[17]	Reputation, Size, Willingness to Customize
[22]	Ability, Intention, Trustee's Promises
[26]	Privacy, Security
[29]	Competence, Motives
[33]	Ability, Integrity, and Benevolence
[36]	Competence, Integrity
[44]	Moral, Integrity, Goodwill
[54]	Anonymity, Security, Transaction Size
[55]	Privacy

In the context of IC, both necessary conditions exist for consumers to trust Internet stores and online transactions. First, the perceived benefits of Internet shopping and/or established customer relations increase the interdependence between customers and online stores. IC is claimed to reduce prices, inventory levels, and the role of brokers [32]. It can reduce the advantages of scale of large retailers, lower the costs of entering international consumer markets, reduce transaction cost, and are more convenient to make a purchase.

Second, Internet shopping involves more uncertainty and risks than traditional shopping. Internet consumers cannot physically check the quality and quantity of the products before receiving a purchase [35]. They cannot monitor the security of sending sensitive personal and financial information, like credit card numbers, through the Internet to a party whose behaviors and motives may be hard to predict [35]. Moreover, the free exchange of electronic information brings the threat of providing easy, and many times unwanted, access to personal information [51]. Consumers cannot control the use of their personal information after they release it to Internet stores during Internet shopping. In sum, the integrity of a store, the security of performing transactions, the accuracy of the transactions, and privacy are the four primary risk factors that cause trust issues to be a concern.

In addition to the analysis of the vulnerabilities that constitute trust, Churchill [14] suggested that extensive literature review and expert opinion provide a sound foundation on which a theoretical domain of trust can be found. Thus, we conducted an extensive review of the literature on trust. The review covered over 120 publications from academic journals, professional magazines, and textbooks, and identified over 40 repeating words that are characteristic of the trustworthiness of Internet stores. We also kept notes on various concerns and vulnerabilities that reflect each of these key words. We listed a sample of studies and their referenced attributes in Table 1. From the list we see that certain attributes like ability and competence are semantically identical or similar. Certain attributes like integrity, security, and privacy appear more frequently than others especially in the context of IC. Some attributes like consistency, atomicity, and discreetness reflect one fundamental concern about accuracy.

To verify the completeness of the list and consolidate conceptual redundancies, we formed a panel of "experts" and asked each member to add overlooked attributes, take away irrelevant ones, and identify similar or identical ones. The panel consisted of three professors respectively in the areas of MIS, Social Psychology, and Organization Studies, four Ph.D. students in Management and Information Systems, and 12 undergraduate students who had experience with Internet shopping. Through a session of brainstorming, we identified a shorter list of 9 items: security, privacy, integrity, accuracy, benevolence, fairness, motives, promise keeping, and capability. By definition, integrity is the perception that an Internet store is honest, ethical, and fair, and adheres to an acceptable set of principles [33]. Thus, integrity captures benevolence, fairness, motives, and promise keeping. Although many studies feel capability is an important dimension of trustworthiness, it is actually an enabler or formative (or causal) factor for other dimensions such as security, privacy, and accuracy. Therefore, the expert panel collectively judged that integrity along with

security, privacy, and accuracy covered the appropriate content domain of trust.

Based on the convergent views on the components of trust, we conceptualize trust as a psychological state comprising the intention to accept vulnerability when shopping in an Internet store. Then we conclude that such vulnerability is reflected in the four aspects: security, privacy, integrity, and accuracy.

Perceived Security: Web providers and users are now sharing the concerns on security that were once raised by environment control [26]. In the context of IC, communication can be intercepted, tampered, and falsified. Personal accounts can be accessed for illegal purposes. Personal properties can be damaged due to malicious attacks and viruses. Consequently, consumers are often unwilling to conduct online transactions for the fear of security breaches and potential damage to personal interests. As a matter of fact, the concern with security has become the top reason for not shopping online [26]. Therefore, to reflect the intention to accept the vulnerability in security, we conceptualize perceived security as the perception that making a transaction with an Internet store is safe. Such a notion of security signals the capability and responsibility of an Internet store in preventing unauthorized data access, illegal data interception, and illegal attacks on customer properties, and in protecting the interests of its customers. Of course, absolute security does not exist. We intend to operationalize the construct in terms of a comparison with traditional means of shopping and whether transmitted data could be intercepted, and accounts broken in.

***Perceived Privacy*:** When associated with consumer activities that take place in the arena of electronic marketplace, the term privacy usually refers to personal information and the protection of its confidentiality. For example, when customers give out their identification numbers, they expect confidentiality that the numbers will only be known by the party who has a legitimate need to know them. Similarly, a customer who buys a bulk of marketing research data may not want his or her competitors to know it; a customer who buys a good of questionable value does not want to disclose his or her identity. The violation of privacy includes unauthorized collection, disclosure, or other misuse of personal information as a direct result of e-commerce transactions. Privacy has been identified as one of the most crucial issues in e-commerce. It is consumers' fear and distrust for potential loss of personal privacy that often makes them unwilling to conduct online transactions [55]. Therefore, to reflect their intention to accept the vulnerability in privacy, we conceptualize perceived privacy as the perception that their personal data with an Internet store are confidential. The concept signals the degree to which an Internet store is capable of and responsible for observing procedural fairness [26] and to exert control on its data collection, access, and secondary use [51]. It reflects a fundamental concern for the loss of proper control on personal data due to improper collection of private information, improper monitoring on Internet activities, and improper transfer of personal information.

***Perceived Integrity*:** In physical commerce, customers often trust a business by its physical locations, facilities, and business licenses. However, they are not as much concerned with these characteristics per se as its integrity to conduct businesses honestly and professionally. The same concern becomes more serious when conducting online transactions. In e-commerce, it is virtually impossible to authenticate the identity of an Internet store through its physical characteristics. Thus, when making online transactions, customers deal with other parties whose true motives and absolute identity are uncertain [54]. Consequently, they are concerned about whether they will receive products or services even though they have paid for them. They are concerned about whether the store will take advantage of them and behave opportunistically. They are also concerned with whether the store has the capacity to offer products and services as it described. To reflect such concerns, we define perceived integrity to be the perception that an Internet store is honest and adheres to an acceptable set of principles [35]. As per the definition, integrity consists of two related aspects of semantics. First, it means that an Internet store does as what it said, i.e., it keeps promises and is procedurally fair. Second, it means that the store says as what it did, i.e., it is honest and credible. In prior studies, the term "reputation" is sometimes considered to be equivalent to perceived integrity. For example, Doney and Cannon [17] defined reputation as the extent to which customers in the industry believe that a company is honest and concerned about its customers. Based on such equivalence, perceived integrity signals the forbearance from opportunism [50] and reflects the capability and responsibility of a store to act professionally.

Perceived Accuracy: In Internet shopping, customers cannot physically touch, check or test a product. All they know about the product is from the descriptions or pictures provided by an Internet store. Besides opportunistic behaviors, customers are also concerned with potential transaction errors such as incorrect product brands, sizes, and quantities, as well as incorrect billing statements. Such errors can occur due to human mistakes. They can also occur due to computer system irregularities such as lack of transaction atomicity [54]. For example, when purchasing a document online, a power failure between sending in payment and obtaining a password to download the document can leave the transaction partially finished. Ideally, such a transaction should be rolled back so that the customer can re-start the process again. However, without atomicity control, the customer will end up with troubles and delays or paying for another transaction. In physical commerce, such errors can be easily corrected using a trip or phone call back to the store. However, making corrections with an online store typically means extra effort and time and sometimes even extra shipping and handling fees. Therefore, Javenpaa et al. [27] suggested that, in order to be able

to trust an Internet store, a customer must believe that the store has both the ability and the motivation to reliably deliver goods and services of the quality expected. Similarly, Butler [6] suggested the dimension of transaction accuracy by emphasizing the criteria such as consistency and discreetness. To capture such a dimension of trust, we conceptualize perceived accuracy as the extent to which a customer believes that transactions with an Internet store are error-free. This concept signals the capability and responsibility of an Internet store in performing its functions accurately.

3. Data Collection

Following the advice by Churchill [14], we utilized the expert panel and a class of undergraduate students to participate in a pre-test and a pilot test respectively. In the pretest, we provided a formal definition of each construct and then a list of measurement items (sentences), which we intend to use to measure the constructs. We asked each member to first read each definition carefully and then give a rating for each item in the 5-point scale to indicate how well the sentence matches the intended construct.

The pretest started with 48 items in total. After the pretest, we analyzed the ratings of each item individually. An item was retained if it was consistently scored 4 or 5 points across the experts. It is dropped if it was consistently rated as no match. For an item that had inconsistent ratings, we adopted alternatives, rephrased it, or dropped it entirely. The pretest substantially refined some of the items by eliminating their ambiguity. The number of items was also reduced to 33.

To further validate the items, we conducted a pilot test using 35 undergraduate students. We randomized the 33 items and created a survey that asked each participant to visit amazon.com and respond to each item by indicating how much he or she agreed with its statement. Then we used the responses and calculated the correlation between each pair of the items. Under each dimensional construct, we retained those items that were highly correlated. However, if an item seemed not to go along with others, we dropped or modified it depending on its content. Through the pilot test, we finally selected 17 items in total for the final test (see Table 2).

The final test involved 173 participants selected from graduate and undergraduate students in two large national universities. Nevertheless, if a subject is not aware of Internet technology and its potential problems, he or she may not make perfect sense of some of the statements in Table 2. Therefore, when identifying participants, we required them to have exposure to electronic commerce. To be representative of the population actually engaging in IC activities, we identified the subjects from students at various stages: 24% from graduate programs, 38% from juniors and seniors, and 38% from freshmen and sophomores. We also distributed the subjects roughly equally in two regions, Midwest and Northeast, in the hope to capture the variation due to urban and rural settings. The subjects were primarily selected from 4 graduate and 8 undergraduate classes on the voluntary basis. Among the individuals who were qualified to participate, the response rate was 73%.

Table 2. Initial Measurement Items for Trust

Perceived Security (5-point Scale Anchored by "Strongly Disagree" and "Strongly Agree")	
PS1	I believe that shopping on this Internet store is just as safe as placing an order by phone
PS2	It is just as safe to make a credit card purchase at this Internet store as it is to make one in person
PS3	The data transmission between my computer and this Internet store is safe
PS4	This Internet store is capable of preventing illegal access
Perceived Privacy	
PP1	I trust that this Internet store will keep customer information confidential
PP2	I am afraid that this internet store might misuse my personal information (Reverse)
PP3	I think it's likely that this store would sell my personal data to others (Reverse)
PP4	This store can be trusted to keep the identities of its customers private
Perceived Integrity	
PI1	I have confidence in this store
PI2	I trust this store to keep my best interests in mind
PI3	I trust this store to act professionally
PI4	Something about this store strikes me as deceptive and misleading (Reverse)
Perceived Accuracy	
PA1	I would trust this store to deliver exactly what I order
PA2	The Internet store can be trusted to fulfill my order accurately
PA3	The store will not overcharge my credit/debit account
PA4	The online product and service information is accurate
PA5	This store will make corrections if my order is in error

In order to ensure the representativeness of our sample to the Internet user population, we collected general and technology demographic data using GVU's WWW User Survey instrument. The participants had ages ranged from 20 to 46 and were on the average 24 years old. 92% of them were English speakers and other 8% spoke Spanish, Chinese, French, etc. More than 90% of the subjects had 2 to 17 years work experience. All except for a few subjects used Internet and web browsers once or several times a day. 72% had purchased goods and services on the Internet and 40% made a purchase for more than 100 dollars. 46% shopped on the Internet frequently. Most of these demographics match the corresponding sample statistics of thousands of Internet users recently surveyed by Georgia Institute of Technology.

Consistent with prior work [27], we used online bookstores to conduct the controlled experiment and to collect data. To avoid possible biases due to the familiarity [21] with a particular store, we searched all online bookstores and identified three at potentially various levels of familiarity to the subjects. We randomly assigned the subjects to the three stores with approximately equal number of subjects assigned to each. Among those assigned to the first

store, all had heard about it and more than half had visited it before. Among those assigned to the second store, 69% had heard about it and 19% had visited it. Among those assigned to the third store, only one had heard about it but none had visited it.

After the participant-store assignment, we provided each subject with a cover page that guides him or her to visit the store, search for a textbook to buy, create an account with the store, and proceed to finish the order. We also provided a credit card number, a billing address, and a social security number for them to create accounts and check out the books. The goal of the experiment was to simulate real Internet shopping experience. After the experiment, each subject was asked to respond to a survey regarding their attitudes to Internet shopping, their perceptions about the store, and their willingness to purchase from the store. For each question, we used a 5-point Likert scale anchored by "Strongly Disagree" and "Strongly Agree." And we gave sufficient time for them to finish the survey.

4. Hypothesized Models of Trust

According to the analysis of the content domain of trust, we see that a common theme of the dimensional constructs is the ability and responsibility of an Internet store. Ability and responsibility form the basis of trust and a lack of either characteristic will lead to no trust. One would not trust an Internet store if it were not capable even though it has good motives. Similarly, he or she would not trust it either if it were not responsible even though it may be capable. On the other hand, if a store is both capable and responsible, what else does one need to be able to trust it? There is possibly none. Therefore, we considered the trustworthiness to be the overall sum of perceived ability and responsibility. Many other authors also suggested the two overall and possibly uncorrelated dimensions of trust. [15, 18] proposed ability and trustworthy intentions as the two overall dimensions of trust. Kee and Knox [29] proposed competence and motives.

Then why do not we measure trust using the two larger dimensions: perceived ability and perceived responsibility? The reason is that they are formative rather than reflective factors of trust. In general, to measure trust using first-order factors, trust must be a common factor underlying the first-order factors rather than be a simple sum of them. The fundamental issue is consistency in directional change among the first-order constructs [11]. In particular, does a directional change in ability imply similar directional shift in responsibility? The answer is possibly negative; an organization is capable does not necessarily imply it is responsible.

On the other hand, ability and responsibility as a common core of trust underlie perceived security, perceived privacy, perceived integrity, and perceived accuracy. If a store is capable of and responsible for its business, it will act professionally to improve its perceived integrity. It will also take a strong measure to ensure security, customer privacy, and transaction accuracy. Therefore, perception in any of the four dimensions is manifested in perceptions in the other dimensions through a larger perception of ability and responsibility. For example, if a customer found that his account had been hacked and modified, he would naturally doubt the ability and/or responsibility of the store and infer that his privacy and expectation of accuracy would be in danger. His perception of its professionalism would be also reduced.

In sum, trust in an Internet store is reflected in the four specific dimensions such as perceived security, perceived privacy, perceived integrity, and perceived accuracy. The overall trustworthiness accounts for the interrelationships among these dimensional factors through the perception of the ability and responsibility of the store. Based on this hypothesis, let us propose and examine six measurement models that are plausible representation of our anticipation.

First-Order Factor Models

Model 1 hypothesizes that one first-order factor — trust in Internet stores — accounts for all the common variance among the 17 items. As we reviewed before, most existing studies have approached trust in Internet stores as a single-dimensional construct [13, 21, 27]. Typical survey questions to assess trust include "This store is trustworthy" [27] and "I trust this store" [21]. If this model is accepted, then it is appropriate to view trust as a single dimensional construct.

Model 2 hypothesizes that two first-order factors account for the variance of 17 items. In this model, perceived privacy and perceived security are combined into one factor, and perceived integrity and perceived accuracy into another. This model emphasizes perceptions in two primary areas: data and goods. The first combination is due to data and is plausible because both security and privacy have a common underpinning: personal information. If data are not secured in transmission and storage, privacy cannot be enforced even though a store does not violate it voluntarily. The second combination reflects that a customer is concerned with whether she will receive goods as expected and be charged correctly regardless whether a discrepancy is due to a mistake or due to a lack of integrity.

Model 3 hypothesizes that 17 items form two first-order factors. However, it combines perceived integrity and perceived privacy into one construct and perceived security and perceived accuracy into another. This model focuses on two overall dimensions: capability and motives. Integrity and privacy are more reflective of the motives of a store whereas security and accuracy are more reflective of its capabilities.

Model 4 hypothesizes still another two first-order factor structure, where perceived integrity is distinct whereas other three are combined into one first-order construct. The justification for this model is as follows. Perceived security, perceived privacy, and perceived accuracy are all

perceptions from the perspective of consumers' own interests such as their security, their privacy, and their orders. On the other hand, perceived integrity is more a perception of the characteristics of an Internet store.

Model 5 hypothesizes that four first-order factors account for the variance of all items. Prior theoretical analysis of content domain provides support for this model. Essentially, this model assumes that every pair of constructs correlate but the correlation is not strong enough to justify for a merger. In addition, this model will act as a benchmark for the test of a second-order model. According to Marsh and Hocevar [39], although a higher-order model is able to explain the covariance of first-order factors, the goodness-of-fit of the higher-order model can never be better than that of the corresponding first-order model. Thus, this model provides a target for testing a second-order model.

A Second-Order Factor Model

Model 6 hypothesizes that four first-order factors account for the variance of the 17 items whereas a second-order factor accounts for the covariance of these first-order factors. Statistically, if Model 5 demonstrates significant interdependence (covariation) among the first-order factors, a natural inference is that there might be a common factor that accounts for the interdependence [42]. If such a common factor exists, then trust will be more than the sum of the four first-order factors. It will consist of the first-order factors as well as the structure of interrelationships among them [51].

Theoretically, ability and responsibility form a common core of trust that underlies the four first-order factors. As we have argued, each first-order construct reflects the ability and responsibility of an Internet store in a specific dimension such as security, privacy, integrity, or accuracy. Therefore, the perceived ability and responsibility, i.e., the trustworthiness of the store, not only extract the variation of these first-order factors but also account for the interrelationships among the first-order factors. Thus, trust in an Internet store may be better measured as the second-order common factor.

5. Data Analysis

To validate the hypothesized models, we employed confirmatory factor analysis using LISREL 8.3 [28]. In the current study, the observed covariance matrix for the 17 measurement items is listed in Appendix. The latent variables include the four first-order factors and one second-order factor. Different measurement models hypothesized in Section 5 underlie different joint distributions of all the variables involved, observed and latent as well. How much each model fits the data can be determined by the extent to which its implied covariance matrix matches the observed one.

As its rationale implies, an important assumption of confirmatory factor analysis is multivariate normality. However, a verification of this assumption is difficult. Instead, we conduct normality test for each observed variable using both normal plots and Kolmogorov-Smirnov statistic. All normal plots show straight lines and all test statistics are strongly significant at $\alpha = 0.000$. They indicate no departure from univariate normality. The test result signals that the multivariate normality holds.

Before conducting confirmatory test on the hypothesized models, we followed the procedure suggested by Segars [47] and tested each first-order factor in isolation first and then in pairs. The procedure can provide the fullest evidence of measurement efficacy and reduce the likelihood of confounds in full structural equation modeling [48]. In the initial phase of isolated model testing, we found that items PA4, PA5, and PS4 have loadings less than 0.6. By analyzing the correlation matrix, we realized that these items seem not to go along with other items under the same construct. Therefore, we deleted these three items to improve the reliability of corresponding constructs. All items under perceived integrity and perceived privacy had reasonably large loadings. However, the modification indices for the test of perceived privacy suggest adding an error covariance between PP2 and PP3. Being reluctant to delete additional items, we add other first-order constructs and conducted paired tests. All test results seem fine except that a similar modification index suggests the existence of error covariance between PP2 and PP3. By analyzing these items, we can see that they both are reversed and both suggest misuse of personal information. Thus, to remove the extraneous correlation that is not captured by the notion of perceived privacy, we deleted PP3. After removing PA4, PA5, PS4, and PP3, all isolated and paired tests went through well.

Finally, we used the remaining 13 items to test the six hypothesized models. Table 3 provides a summary of the model-fit indices and their thresholds recommended by previous studies. As shown, by all measures of fit Models 5 (four first-order factors) and 6 (the second-order factor model) are deemed excellent while alternative models being deemed unacceptable. Models 5 and 6 both had insignificant χ^2 statistics with p-values above 0.4, which is far higher than the threshold 0.05. They satisfied other cited criteria in terms of GFI, NFI, RMR, RMSEA, and the ratio of χ^2 to degree of freedom, as well as the more stringent criterion of AGFI [12, 23]. Most strikingly, their corresponding probabilities of close fit, i.e., RMSEA < 0.05, are all higher than 0.95. It means that the type I error of rejecting a not-close fit hypothesis is less than 0.05 [5].

It is interesting to observe from Table 3 that, while all often-cited fit indices being able to tell a good model fit from a bad one, their powers of detection are not equal. According to the ratio of χ^2 to degree of freedom, Models 2 and 3 might be declared acceptable because their ratios are close to 2. Similarly, the NFI, GFI, and AGFI for Models 1-4 are all close to 0.8 and some values are even close to 0.9. The corresponding values of RMR are close to or less than 0.08. Using less stringent criteria, these models would

be judged acceptable. However, by using the p-value of χ^2 statistic and P(RMSEA < 0.05), these models fall apart and a clear distinction from Models 5 and 6 can be identified.

Table 3: Measures of Model Fit: Alternative Models

Model	**1**	**2**	**3**	**4**	**5**	**6**	
c^2	228.61	142.19	146.20	196.40	57.24	63.13	
df	65	64	64	64	59	61	
c^2/df	3.517	2.222	2.284	3.069	0.970	1.035	<2.00
c^2 Sig.	0.000*	0.000*	0.000*	0.000*	0.541	0.401	>0.05
NFI	0.785	0.825	0.819	0.818	0.935	0.927	>0.90
CFI	0.841	0.896	0.890	0.876	0.998	0.992	
GFI	0.830	0.887	0.884	0.851	0.951	0.947	>0.90
AGFI	0.762	0.840	0.836	0.788	0.925	0.920	>0.90
NNFI	0.809	0.874	0.866	0.849	0.997	0.990	
RMR	0.080	0.068	0.075	0.074	0.041	0.046	<0.05
RMSEA	0.121	0.084	0.086	0.110	0.000	0.014	<0.05
P(<.05)	0.000	0.002	0.001	0.000	0.977	0.956	>0.90

Table 4. The Indices of Convergent Validity

	PS	PP	PI	PA
Composite Reliability	0.79	0.76	0.83	0.71
Cronbach a	0.76	0.80	0.83	0.71
AVE by Regressions	0.56	0.51	0.55	0.45
TVE by Principal Factor	0.68	0.72	0.66	0.63

Table 5. Results of Discriminant Validity Tests

Test	Original χ^2	Alternative χ^2	χ^2 Difference
PI			
PA	10.56 (13)	42.28 (14)	31.72***
PS	9.59 (13)	74.06 (14)	64.47***
PP	11.70 (13)	81.45 (14)	69.75***
PA			
PS	14.10 (8)	54.23 (9)	40.13***
PP	5.31 (8)	72.03 (9)	66.72***
PS			
PP	6.49 (8)	76.05 (9)	69.56***

Convergent Validity

Figure 1 illustrates the structure and estimated parameters of the four-construct, first-order factor model (Model 5). As shown, the indicator loadings of items to their respective constructs are all above 0.60, indicating that each measure is accounting for 50 percent or more of the variance of the underlying latent variable [11]. The t-values obtained for the coefficients range from 8.317 to 13.564, indicating that all factor loadings are significant at the level $\alpha = 0.0001$. The significance level is far in excess of the critical value 0.01 suggested by Hair et al. [23]. Both loadings and their significance levels provide strong evidence to support the convergent validity of the items [1].

The composite reliability indices of [19] are listed in Table 4. As shown, they are all in excess of 0.70, implying acceptable level of reliability for each of the constructs [28]. As a comparison, we also show the corresponding Cronbach α coefficients, which are also higher than the acceptable threshold 0.7 [41]. As a similar indicator of measurement reliability, average variance extracted (AVE) represents how much variance in each item on the average is explained by the corresponding construct [19]. It is conceptually similar to the total variance extracted (TVE) in principal factor analysis. The AVE values based on the formula of [19] are listed in Table 4. As shown, except for perceived accuracy, the AVE for each construct is above 0.5. It indicates that, on the per-item average basis, the amount of variance captured by the first-order construct is more than the amount of variance due to measurement error. The AVE of perceived accuracy is 0.45, which is a little bit lower than 0.50 (see Section 7 for a discussion). As a comparison, we computed the TVE value for each construct by using the eigenvalues of the correlation matrix of its scale items. As shown, all TVE values are higher than 0.6, indicating that more than 60% of total variance contained in all the items is captured by the corresponding factor. In sum, all indices suggest the first-order constructs exhibit strong properties of convergent validity.

Discriminant Validity

Discriminant validity is the extent to which items measuring two distinct constructs are shown to be empirically distinct and not so highly related to each other [7]. To do the test, we made a series of comparisons between the original model, where two constructs are treated as distinct, and the alternative model, where they are united as one construct. Discriminant validity is implied if the χ^2 statistic of the original model is significantly lower than that for the alternative model; this suggests that the original model has a better model fit.

Table 5 shows the results in all six paired comparisons. All the χ^2 differences are highly significant at $p < 0.001$. Hence, each item seems to capture a construct that is significantly unique from other constructs, providing strong evidence of discriminant validity. Also importantly, the estimated correlation between each pair of constructs is below the suggested cutoff value 0.90 [19], indicating distinctness in construct content.

Testing the Second-Order Model

Both convergent and discriminant validities indicate how well the first-order constructs are defined and measured. However, our eventual goal is to determine how well trust as a higher-order construct captures the variance and covariance of these first-order constructs. To formally test the validity of the second-order factor model (Model 6), we need to first compare its model fit with that of the baseline model (Model 5). Compared to the baseline model, the second-order factor model explains the covariance among first-order factors in a more parsimonious way. Thus, even when the higher-order model is able to explain the factor covariance, its goodness-of-fit can never be better than the corresponding first-order model. In this sense, the first-order model provides an optimum fit or target for the higher-order model [39]. It has been suggested that the

efficacy of a second-order model be assessed using the so-called target coefficient, i.e., the ratio of χ^2 (baseline model) to χ^2 (second-order model). This coefficient has an upper bound of 1.0 with higher values indicating the higher power of the second-order factor in capturing the covariance among first-order factors. Figure 2 shows the structure and estimated parameters of the second-order factor model of trust (Model 6). The overall χ^2 is 63.13 that is insignificant with a p-value = 0.40 (see Table 3). Adjusting the degree of freedom, the normed value of χ^2 is 1.04, indicating an excellent model fit and no evidence of over-fitting. The target coefficient is a very high value 0.91, indicating that the introduction of the second-order factor into the baseline model does not significantly increase χ^2. Since the second-order model is more parsimonious, it should be accepted as a better representation of the "true" factor structure according to Occam's razor [42].

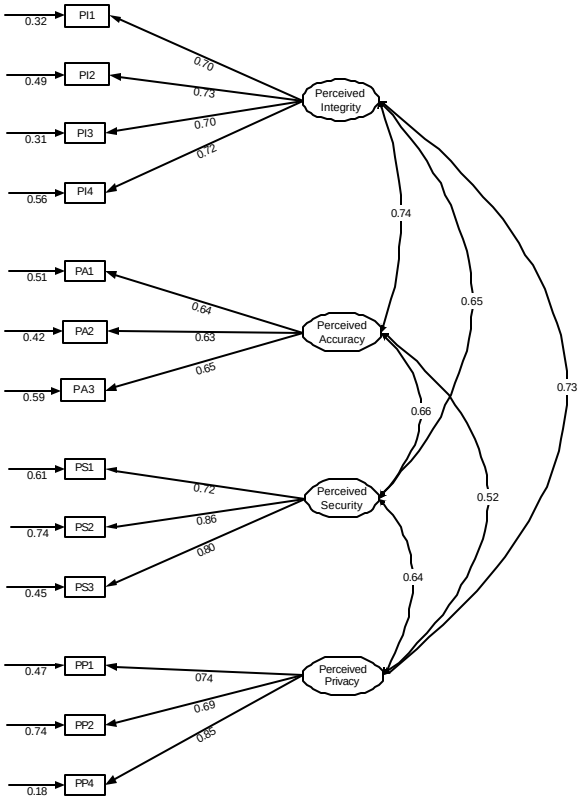

Figure 1. A First-Order Model of Trust

In addition to target coefficient, the power of the second-order factor in explaining the variance of the first-order factors provides another piece of evidence in support of the second-order factor structure. The path loadings to perceived security, perceived privacy, perceived integrity, and perceived accuracy are respectively 0.77, 0.78, 0.92, and 0.78 with t-values raging from 7.03 to 9.9. The corresponding R^2 values are respectively 0.59, 0.61, 0.85, and 0.61. Similar to the loadings from a first-order factor to its observed items, these loadings and their significance show a strong convergent validity and reliability of the second-order factor [11]. To obtain overall validity indices, we can similarly compute the composite reliability and AVE of the second-order construct, which are respectively 0.89 and 0.66. Segars and Grover [48] noted that the most convincing evidence of the explanation power is the observed total coefficient of determination. This statistic is 0.94 for Model 6, indicating that a large amount of variance and covariance among the four first-order constructs is explained by the second-order factor and captured by the regression models.

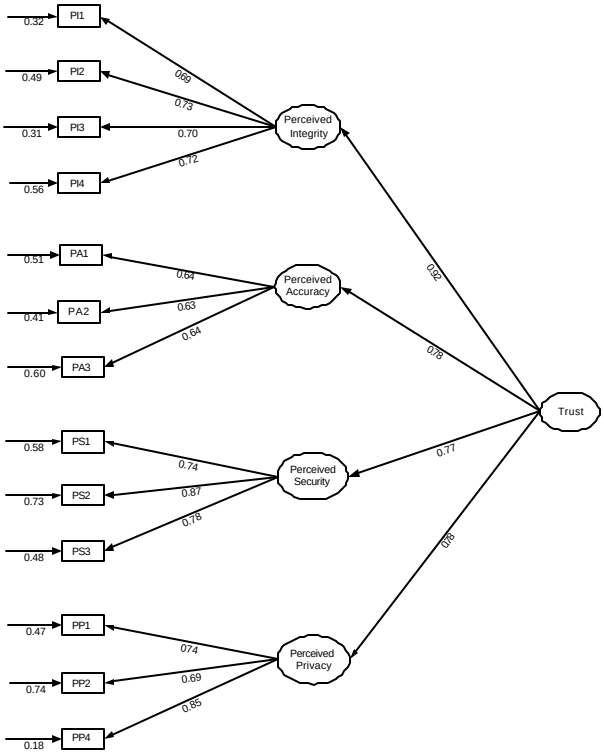

Figure 2. A Second-Order Factor Model of Trust

All evidence in support of the second-order model has been based on type I errors. Then, is it likely that we accept a wrong hypothesis of close fit but in fact the model has a bad or mediocre fit? To answer this question, we conducted power analysis for tests of fit using the technique proposed in [27, 37] and considered RMSEA < 0.08 to be a good fit. Browne and Cudeck [5] suggested a more restrictive criterion: RMSEA < 0.05 for close fit and RMSEA between 0.08 and 0.1 for mediocre fit. Based on the latter stringent criterion, we used the SAS program provided by MacCallum et al. [37] and computed the power indices with various RMSEA values in [0.08, 0.1] representing mediocre fit. We found that Model 6 has a power between 0.77 and 0.99, implying that more than 77% of time we can reject a hypothesis of close-fit if the model indeed has a mediocre fit. By using a more lenient criterion, such a power can increase to almost 100%. Therefore, based on Type II errors, the second-order model will be still considered to fit data well.

6. Conclusions and Discussions

From the perspective of potential risks and vulnerabilities involved in online shopping, we conceptualized the first-order constructs of perceived security, perceived privacy, perceived integrity, and perceived accuracy, and proposed measuring trust in an Internet store using these dimensions. Then, by conducting an extensive review of literature on trust, a pre-test, and a pilot test, we operationally defined the first-order constructs and developed scale items to measure them. We conducted a controlled experiment that provided a simulated online shopping experience to all participants before they responded to our survey.

To be consistent with a wide range of existing studies on trust, we believed ability and responsibility to be an overall dimension that governs how the trustworthiness of an Internet store is perceived. Therefore, we hypothesized the equation that trust = perceived ability + perceived responsibility. Then, based on this equation, we theoretically justified that the first-order factors —perceived security, perceived privacy, perceived integrity, and perceived accuracy—are the reflections of a single higher-order construct, trust. Each first-order factor is manifested in other first-order factors through trust. In addition to this second-order factor model, we proposed 5 alternative first-order factor models, with a four-factor model as the target for benchmarking the test of the second-order model.

We employed confirmatory factor analysis using LISREL. We used selected items and tested the six hypothesized models. We reported all model fit indices cited in the existing information systems research and additional ones such as the significance of χ^2 and the p-value for RMSEA < 0.05, which we felt are powerful in separating a good model from mediocre ones. All the indices and their corresponding criteria clearly indicate the superiority of the second-order factor model and the four first-order factor model while rejecting the acceptance of other alternative models. By using the statistics of the first-order model, we determined that the four first-order factors are empirically valid in terms of their convergent and discriminant validities. Finally, we determined that the second-order model is a better representation of the factor structure than the first-order counterpart based on its parsimony, the target coefficient, and the total coefficient of determination. We also empirically determined the power of such a test of the second-order model to be close or higher than 0.8.

Before we discuss the implications of this study, its limitations should be noted. First, the use of student subjects and Internet bookstores may limit the generalizability of the results. Although we carefully simulated and controlled many parameters so that our sample is representative of the Internet user population, a further replication using real online customers might be worthwhile. The second limitation is about the use of confirmatory factor analysis. As we noted, the technique essentially validates a joint distribution assumption using the conformance of its marginal to observed data. There is a possibility that there exist equivalent models [24]. In this study, we build our models based on a conceptual foundation. We also explored alternative models. The chance of having equivalent models is slim. Third, our measurement model has an excellent model fit. In comparison, the reliability of perceived accuracy is a bit low. Its AVE is 0.45, which is below the recommended value 0.5 [19]. The problem is largely due to the early stage of research in the area. We found a similar problem in some existing studies.

All statistical evidence converges and is in support of our conceptualization that trust is a multi-dimensional construct and is well measured by perceived security, perceived privacy, perceived integrity, and perceived accuracy. We also found that the notion of trust, as a second-order construct, accounted for most of variance and covariance in the first-order factors. These results have several implications. First, they imply that trust is more complicated than previously thought. In general, consumers manifest their trust in an Internet store through their perceptions in security, privacy, integrity, and accuracy. In other words, customers reflect their trust in an Internet store through their perceptions that their shopping activities are safe, their privacy is protected, their transactions are error-free, and the store acts professionally. Furthermore, the results indicate that consumers have vulnerability concerns in all these aspects rather than in any particular dimension and that the interrelationships among these factors are an important component of accurately measuring trust. Second, the results imply that the measurement of trust in an Internet store, although complicated, can be done through an indirect measurement of the first-order factors. Such a measurement model will provide an important metric of the effectiveness of an Internet store for its managers and/or an important metric of vulnerability, perceived risk, as well as expectations of consumers when shopping online. Third, since each first-order construct is manifested in the others through perceived ability and responsibility, our results imply that a store can manipulate certain variables in order to improve consumer trust. For example, by providing accurate billing statements or filling customer orders accurately, a store can convey a sense of ability and responsibility. Such a sense will improve the customer perception in security and integrity that will otherwise be difficult to achieve.

References

[1] Anderson, J.C., & Gerbing, S.W. 1988. Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin* 103(3) 411-423.

[2] Benbasat, I. 1991. Commentary. K. L. Kraemer, ed. *The Information Systems Research Challenge: Survey Research Methods* vol. 3. Harvard Business School, Boston, MA.

[3] Bhattacharya R., Devinney T. M., & Pillutla M. M. 1998. A formal model of trust based on outcomes. *Academy of Management Review* 23(3) 459-472.

[4] Brewer D. 1999. Keeping virtual worlds open for business. *Telecommunications* 33 65-66.

[5] Browne, M. W. & Cudeck, R. 1993. Alternative ways of assessing model fit. In Bollen, K. A. & Long, J. S. (Eds.). Testing structure equation models. Newbury Park, CA. Sage. 136-162.

[6] Butler, J. K. Jr. 1991.Toward understanding and measuring conditions of trust: evolution of conditions of trust inventory. *Journal of Management* 17(3) 643-663.

[7] Campbell, D. T. & Fisk D. W. 1959. Convergent and discriminant validation by the multitraite – multimethod matrix. *Psychological Bulletin* 16 81-105.

[8] Castelfranchi, C., & Tan, Y. H. (eds.) 2002. Trust and deception in virtual societies. Dordrecht, Netherlands: Klumer Academic Publishers, forthcoming.

[9] Cheung, C. & Lee, M. K.O. 2001. Trust in Internet shopping: Instrument development and validation through classical and modern approaches. *Journal of Global Information Management* 9,23-35.

[10] Chiles, T. H., & McMackin, J. F. 1996. Integrity variable risk preference, trust, and transaction cost economics. *Academy of Management Review* 21 73-99.

[11] Chin, W.W., 1998. Issues and opinions on structural equation modeling. *MIS Quarterly.*

[12] Chin, W.W. and Todd, P.A. 1995. On the use usefulness, and ease of use of structural equation modeling in MIS research: A note of caution. *MIS Quarterly*.19(2) 237-245.

[13] Chircu, A. M., Davis, G.B. & Kauffman, K. J. 2000. The role of trust and expertise in the adoption of electronic commerce intermediaries. *Working paper*.

[14] Churchill, G. A. 1979. A paradigm for developing better measures of marketing constructs. *J. of Marketing Research* 16 64-73.

[15] Cook, J. & Wall, T. 1980. New work attitude measures of trust, organizational commitment, and personal need nonfulfillment. *Journal of Occupational Psychology* 39-52.

[16] Coleman, J. S., 1990. *Foundations of Social Theory.* Cambridge, MA: Belknap Press.

[17] Doney, P. M., & Cannon, J. P. 1997. An examination of the nature of trust in buyer-seller relationships. *Journal of Marketing* 61(2) 35-51.

[18] Deutsch, M. 1960. The effect of motivational orientation upon trust and suspicion. *Human Relations* 13 123-140.

[19] Fornell, C. & Larcker, D. F. 1981. Evaluating structural equation models with unobserved variables and measurement error, *Journal of Marketing Research* 39-50.

[20] Gambetta, D. 1990. Can we trust Trust?, In *Trust: Making and Breaking Cooperative Relations*, Gambetta, D. (ed.), Basil Blackwell, Oxford 213-237.

[21] Gefen, D. 2000. E-commerce: The role of familiarity and trust. *Working paper*.

[22] Good, D. 1988. Individuals, interpersonal relations, and trust. In Gambetta, D. G. (ed.). Trust: Making and breaking cooperative relations. Oxford: Basil Blackwell. 31-48.

[23] Hair, J. F. Jr., Anderson, R. E., Tatham, R. L. & Black, W. C. 1998. *Multivariate Data Analysis with Readings*, 5th Edition. Englewood Cliffs, NJ: Prentice Hall.

[24] Hayduck, L. A. 1996. *LISREL Issues, Databases, and Strategies*. The Johns Hopkins University Press, Baltimore and London.

[25] Hinkin, T. R. 1998. A brief tutorial on the development of measures for use in survey questionnaires. *Organizational research methods* 16(1) 104-121.

[26] Hoffman, D. L., Novak, T. P. & Peralta M. 1999. Building consumer trust online. *Communications of the ACM* 42(4) 80-85.

[27] Javenpaa, S. L., Tractinsky N., & Vitale M. 2000. Consumer trust in an Internet store. *Information Technology and Management* 45-71.

[28] Jöreskog, K.G., & Sörbom, D. 1989. LISREL 7 user's reference guide. *Scientific software*, Chicago, IL.

[29] Kee, H. W. & Knox, R. E. 1970. Conceptual and methodological considerations in the study of trust. *J. of Conflict Resolution*. 14 357-366.

[30] Keen, P .G. W. 1997. Are you ready for "trust" economy. *ComputerWorld* 31(16) 80.

[31] Keeney, R. L. 1999. The value of Internet commerce to the customer. *Management Science* 45(4) 533-542.

[32] Lee, H. G. 1998. Do electronic marketplaces lower the price of goods. *Communications of the ACM* 73-80.

[33] Lee, M. K. O. & Turban, E. 2001. A trust model for consumer Internet shopping. *International Journal of Electronic Commerce* 6(1) 75-91.

[34] Lewicki, R. J. & McAllister, D. J. 1998. Trust and distrust: measuring the difference. *Unpublished Manuscript*, Ohio State University.

[35] Lewis, J. D. & Weigert, A. 1985. Trust as a social reality. *Social Forces* 63 967-985.

[36] Lieberman, J. K. 1981. *The Litigious Society.* New York: basic Books.

[37] MacCallum, R .C., Browne, W. & Sagawara, H. M. 1996. Power analysis and determination of sample size for covariance structure modeling. *Psychology Method* 130-149.

[38] MacCrimmon, K. R. & Wehrung, D. A. 1986. *Taking Risks: the Management of Uncertainty*. New York: Free Press.

[39] Marsh, H. W. & Hocevar, D., 1985. Application of confirmatory factor analysis to the study of self-concept: First and higher order factor models and their invariance across groups. *Psychological Bulletin* 97(3) 562-582.

[40] Mayer, R. C., Davis, J. H. & Schoorman, D. 1995. An integrative model of organizational trust. *Academy of Management Review* 20 709-734.

[41] Nunnally, J. C. 1967. *Psychometric Theory*. New York: McGraw-Hill.

421] Pearl, J. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, UK.

[43] Peter, J. P. 1981. Construct validity: review of basic issues and marketing practices. *Journal of Marketing Research* 18(2) 133-145.

[44] Ring, S. M. & Van de Ven, A. 1992. Structuring cooperating relationships between organizations. *Strategic Management Journal* 13 483-493.

[45] Rotter, J. B. 1967. A new scale for the measurement of interpersonal trust. *Journal of Personality* 35 615-665.

[46] Rousseau D. M., Sitkin, S. B., Burt, R.S. & Camerer, C. 1998. Not so different after all: a cross-discipline view of trust. *Academy of Management Review* 23(3) 393-404.

[47] Segars, A.H. 1997. Assessing the unidimensionality of measurement scales: A paradigm and illustration within the context of information systems research. *Omega* 25(1) 107-121.

[48] Segars, AH. and Grover V. 1998. Strategic information systems planning success: An investigation of the construct and its measurement. *MIS Quarterly* 139-163.

[49] Senge, P. M. 1990. *The Fifth Discipline.* Currency Doubleday, New York.

[50] Smith, J. B. & Barclay, D. W. 1997. The effects of organizational differences and trust on the effectiveness of selling partner relationships. *Journal of Marketing* 61 3-21.

[51] Stewart, K. A. & Segars, A. H. 2002. An empirical examination of the concern for information privacy instrument. *Information Systems Research* 13(1) 36-49.

[52] Straud D. W., Hoffman D. L., Weber B. W. & Steinfield C. 2002. Measuring e-commerce in net-enabled organizations. *Information Systems Research* 13(2) 115-124.

[53] Straud, D. W. & Watson, R. T. 2001. Research commentary: Transformational issues in researching IS and net-enabled organizations. *Information Systems Research* 12(4) 337-345.

[54] Tygar, J. D. 1998. Atomicity in electronic commerce. *Mixed Media* 32-43.

[55] Wang, H. Q., Lee, M. K. O. & Wang, C. 1998. Consumer privacy concerns about internet marketing. *Communications of the ACM* 41 63-70.

[56] Watson, R. T., Akselsen, S. & Pitt, L. F. 1998. Attractors: building mountains in the flat landscape of the world wide web. *California Management Review* 40(2) 36-56.

[57] Williamson, O. E. 1993. Calculativeness, trust and economic organization. *Journal of Law and Economics* 30 131-145.

[58] Zmud, R. W. & Boynton, A. C. 1991. Survey measures and instruments in MIS: Inventory and appraisal. K. L. Kraemer, ed. *The Information Systems Research Challenge: Survey Research Methods*, vol. 3, Harvard Business School, Boston, MA.