
Individual Trust in Online Firms: Scale Development and Initial Test

ANOL BHATTACHERJEE

ANOL BHATTACHERJEE is an Associate Professor of Information Systems at the University of South Florida. He received a Ph.D. and an M.B.A. from the University of Houston and prior B.S. and M.S. from Indian Institute of Technology (India). His research interests include behavioral issues in electronic commerce, information technology (IT) acceptance and usage, IT-enabled organizational transformation, and IT-driven knowledge management. His prior research has been published in *Information Systems Research*, *Journal of Management Information Systems*, *MIS Quarterly*, *Decision Sciences*, *Decision Support Systems*, *IEEE Transactions on Systems, Man, and Cybernetics*, *Data Base, Information & Management*, and several other refereed journals.

ABSTRACT: The importance of trust as a key facilitator of electronic commerce is increasingly being recognized in academic and practitioner communities. However, empirical research in this area has been beset by conflicting conceptualizations of the trust construct, inadequate attention to its underlying dimensions, causes, and effects, and lack of a validated trust scale. This paper addresses these limitations in part by theoretically conceptualizing and empirically validating a scale to measure individual trust in online firms. The proposed scale taps into three key dimensions of trust: trustee's ability, benevolence, and integrity. An iterative testing and refinement procedure using two field surveys of online retailing and online banking users, leads to a final seven-item trust scale that exhibits adequate levels of reliability, convergent validity, discriminant validity, and nomological validity. It is expected that the scale presented in this paper will assist future empirical research on trust in online entities.

KEY WORDS AND PHRASES: e-commerce metrics, electronic commerce, online trust, scale development, trust.

THE CALL FOR THEORETICALLY AND EMPIRICALLY RIGOROUS instruments to measure unobservable constructs has been echoed time and again by information systems (IS) researchers (for example [5, 42, 46]). Such calls have inspired the development and validation of instruments to measure perceived usefulness and ease of use [13], micro-computer playfulness [48], computer self-efficacy [9], and strategic information planning success [43]. Instrument construction plays an important role in advancing research by bringing rigor to the process of scientific inquiry, improving the replicability and trustworthiness of research findings, and permitting confirmatory

research [46]. One area that has attracted significant interest among IS scholars (for example [18, 23, 33, 34]), but has received little attention to scale development, is trust in electronic commerce. This paper presents a theoretically derived scale for measuring individual trust in online firms and validates it iteratively using multiple field surveys.

Why is trust important in electronic commerce? Trust is an important component of online exchange relationships characterized by uncertainty, anonymity, lack of control, and potential opportunism [21]. Online transactions often require sharing of sensitive personal information (such as, mailing address, telephone number), corporate information (such as inventory data), and financial information (such as credit card numbers) among the transacting parties. Trust helps assure that one party will not take advantage of the vulnerability of the other during or after the transaction. Trust acts as an informal control mechanism that reduces friction, limits opportunistic behaviors, minimizes the need for bureaucratic structures, encourages future transactions, and helps build long-term relationships. In contrast, lack of trust may require full disclosure of otherwise commonly accepted transaction rules in the form of extensive legal contracts, leading to greater transaction costs and lower business efficiencies [11]. Hence, initiating, building, and maintaining trust between buyers, sellers, and partners are widely believed to be the key drivers of success for most online firms [25].

However, developing trust in online relationships is substantially more difficult than in traditional (offline) relationships [8, 26]. In offline ("brick and mortar") commerce, trust typically results from a firm's physical proximity, history of successful prior transactions, and legal restrictions motivating "appropriate" business conduct. In contrast, online firms are often physically located in a different part of the country, have limited history of prior online transactions, and legal policies governing e-commerce are not yet established or even well understood. Moreover, many online firms view themselves as intermediaries (such as, auctioneers, catalog aggregators) and are unwilling to be held accountable for individuals or firms misappropriating their infrastructure. As a result, 53 percent of online users do not trust commercial Web sites collecting personal data, 66 percent do not register on online sites in fear that their personal data may be used inappropriately, and 40 percent falsify information when registering online [19]. As stated by Hoffman et al., "consumers simply do not trust most Web providers enough to engage in 'relationship exchanges' involving money and personal information" [21, p. 80]. If not redressed, lack of trust may eventually pose a substantial threat to the successful conduct of e-commerce [28, 33, 37, 45].

Academic research on trust as a driver of e-commerce is beginning to emerge (for example [18, 23, 33, 34]). However, trust research is beset by conflicting conceptualizations of the trust construct [32], inadequate understanding of the relationships between trust, its antecedents, and consequents [40], and current trust scales are neither theoretically derived, nor rigorously validated. Further, prior trust scales were not specifically designed for use in online settings and may not be entirely appropriate in such settings. This paper addresses these problems by conceptualizing, operationalizing, and validating a theoretical scale for measuring individual trust in online firms. An initial ten-item trust scale is derived based on theoretical dimensions of trust pro-

posed in the literature, which is validated iteratively via two field surveys in online retailing and online banking contexts and refined to a seven-item scale. The final scale demonstrates adequate psychometric properties, including reliability, convergent and discriminant validity, and nomological validity.

Theoretical Foundations

Conceptualizing Trust

TRUST, AS A RESEARCH CONCEPT, has been examined in several social science disciplines, including sociology, social psychology, anthropology, economics, marketing, organizational behavior, and most recently, e-commerce. Synthesizing disciplinary perspectives, Mayer et al. proposed an integrative definition of trust as "the willingness of a party [trustor] to be vulnerable to the actions of another party [trustee] based on the expectation that the other [trustee] will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party [trustee]" [32, p. 712]. In consumer-based e-commerce, consumers (users) are the trustors and online firms are the trustees, since users provide sensitive information such as e-mail addresses, credit card numbers, and personal preferences to these firms and are hence vulnerable to firm behavior. Users have limited ability to monitor or control firms' use of their private information; hence, the need for trust.

Certain common threads can be abstracted from the multiplicity of trust conceptualizations across disciplines [32, 40]. First, individual level trust (such as in personal or work relationships) is distinct from that at group level (such as interdepartmental) or firm level (such as in supply chains) [10]. For instance, one's trust in a selling firm is different from one's trust in the salesperson conducting that sale or the social system within which the sale was conducted [49]. Hence, trust scales intended to assess user trust in individual or group settings may not be directly applicable to online firms.

Second, trust can be viewed as a general personality trait or a domain-specific psychological state [32]. Personality traits are relatively stable intrinsic characteristics shaped by developmental and social factors extraneous to a given context and are invariant to situational stimuli [48]. In contrast, psychological states are affective or cognitive episodes that fluctuate with situational contexts and may be influenced by the person's interaction with a situation. Organizational research indicates that overall personality traits are less predictive of specific behaviors than domain-specific states, because of their inability to distinguish between situational differences [15, 30, 32]. Hence, the personality trait characterization of trust may not adequately explain why an individual who is relatively trusting in personal life may demonstrate less trust toward anonymous Web retailers. Based on these contentions, trust is viewed in this study as a domain-specific psychological state.

Third, trust has been conceptualized variously as a belief, attitude, intention, and behavior [32, 34]. As a psychological state, trust is clearly distinct from, but antecedent to, behavior. Mayer et al.'s [32] definition of trust as willingness to take risks

seems to view trust as an intention; however, their underlying dimensions of trust as well as those of other trust theorists (for example [38]) portray trust as a collection of beliefs (such as, benevolence, honesty, confidence, and so on). Trust beliefs refer to trustor's perceptions of trustee attributes that may influence trustee's behavior. Although attitude reflects human affect, beliefs may include both cognitive and affective components, and hence, trust beliefs represent a more holistic conception of trust. Hence, trust is viewed here as an aggregation of beliefs.

Fourth, prior literature reveals several related but distinct dimensions (beliefs) of trust, including benevolence, honesty, faith, dependability, predictability, fairness, honesty, confidence, security, hope, openness, integrity, promise fulfillment, being considerate, and many others (for example [23, 32, 38, 40, 49]). Despite semantic overlaps across these dimensions and greater salience of some dimensions in specific contexts, it is commonly agreed that a multidimensional view of trust can best capture the breadth and complexity of this complex construct. The next subsection discusses trust dimensions salient to online contexts.

Dimensions of Trust

In his study of the formation and evolution of effective corporate relationships, Gabarro [17] identified nine "bases" of trust: integrity, motives, consistency, openness, discreetness, functional competence, interpersonal competence, business sense, and judgment. Butler [7] expanded on Gabarro's work by interviewing 84 corporate managers to identify 10 trust categories: availability, competence, consistency, discreetness, fairness, integrity, loyalty, openness, promise fulfillment, and receptivity. Among other studies, Larzelre and Huston [27] described interpersonal trust as consisting of benevolence and honesty; Rempel et al. [38] defined trust in spousal relationships in terms of predictability, dependability, and faith; Cummings and Bromiley [11] conceptualized interdepartmental trust as keeping commitments, negotiating honestly, and avoiding taking excessive advantage; Doney and Cannon [14] viewed credibility and benevolence as components of trust in supply chain partners; and Zaheer et al. [49] found predictability, reliability, and fairness as salient dimensions of trust in salesperson and selling firms. Table 1 presents a cross-disciplinary sampling of 15 conceptual and empirical studies of trust, along with their context, type, and proposed dimensions.

In order to make sense of the above multitude of trust dimensions and explicate dimensions relevant to online settings, it is first useful to reconcile the proposed dimensions into a manageable few. To that end, first, it can be debated whether some of the proposed dimensions represent trust at all. For instance, Luhmann [31] distinguished confidence from trust by suggesting that trust inherently entails the assumption of risk (a probability of loss) since absence of risk obviates the need for investing in trust-building initiatives; however, confidence eliminates risk by ignoring possible alternatives and is therefore distinct from trust. Mayer et al. [32] differentiated between cooperation and trust by arguing that, although trust can lead to cooperative behavior, trust is not necessary for cooperation to occur (since cooperation does not

Table 1. Trust Dimensions from Prior Literature

Study	Context	Type/Goal of study	Trust dimensions
Butler [7]	Interpersonal trust among corporate managers	Empirical: Developing an inventory of trust conditions	Availability, competence, consistency, discreetness, fairness, integrity, loyalty, openness, promise fulfillment, receptivity
Couch et al. [10]	Interpersonal trust in romantic, peer, and generalized others	Empirical: Developing a 40-item interpersonal trust scale	Unspecified (collection of items from prior scales)
Cummings and Bromiley [11] Doney and Cannon [14]	Inter-unit or interdepartmental trust Trust in salesperson and supplier firm in buyer-seller relationships	Empirical: Developing a 62-item scale to measure interdepartmental trust Empirical: Testing a model of trust in supplier firm and salesperson	Keeps commitments, negotiates honestly, avoids taking excessive advantage Credibility, benevolence
Gabarro [17]	Interpersonal trust between corporate executives	Empirical: Understanding relationship formation among senior executives	Integrity, motives, consistency, openness, discreetness, functional competence, interpersonal competence, business sense, judgment
Hart et al. [20]	Inter-unit trust at General Motors	Empirical: Understanding dimensions of organizational trust	Openness/congruity, shared values, autonomy
Jarvenpaa et al. [22]	Interpersonal trust in virtual team members	Empirical: Testing a trust model in virtual team setting	Ability, benevolence, integrity
Jarvenpaa et al. [23]	Consumer trust in Internet store	Empirical: Testing a trust model for online shopping behavior	Unspecified (trust scale taken from Doney and Cannon [14])

(continued)

Table 1. Trust Dimensions from Prior Literature (continued)

Study	Context	Type/Goal of study	Trust dimensions
Johnson-George and Swap [24]	Interpersonal trust	Empirical: Developing a nine-item interpersonal trust scale	Emotion, reliableness
Larzelre and Huston [27]	Trust in close interpersonal relationships	Empirical: Developing an eight-item dyadic trust scale	Benevolence, honesty
Lewicki et al. [29]	Unspecified	Conceptual	Hope, faith, confidence, assurance, initiative
Mayer et al. [32]	Unspecified	Conceptual	Ability, benevolence, integrity
McKnight et al. [34]	Individual trust in organizations	Conceptual: Explaining paradox of high initial trust	Unspecified (benevolence, competence, honesty, and predictability as antecedents of trust)
Rempel et al. [38]	Trust in spousal relationships	Empirical: Developing a 26-item trust scale	Predictability, dependability, faith
Sheppard and Sherman [44]	Unspecified	Conceptual	Discretion, reliability, competence, integrity, concern, predictability, consistency, empathy
Zaheer et al. [49]	Interorganizational and interpersonal trust in interfirm relationships	Empirical: Understanding how trust affects interorganizational exchange	Reliability, predictability, fairness

necessarily put either party at risk) and one may cooperate for reasons other than trust. They also distinguished predictability from trust by holding that, although prediction and trust are both means of uncertainty reduction, trust must go beyond predictability to embrace vulnerability in a relationship. Likewise, some of the other proposed dimensions, such as autonomy, emotion, hope, and motives, are also unrelated to trust by virtue of lacking the risk component.

Second, many trust dimensions were originally intended for interpersonal (that is, family or work) relationships and may not be salient for online firms. For instance, availability and accessibility relate to interpersonal trust because their absence introduces risk in interpersonal relationships. However, the same dimensions are less relevant in online settings since firm interfaces are typically available around-the-clock and accessible from any Internet-connected computer. In contrast, some individual-level attributes of trust (such as, altruism, receptivity, discretion) may be extended to e-commerce contexts, since firms, as collections of humans, can be viewed as demonstrating anthropomorphic (human-like) behaviors [49].

Following an extensive review of the trust literature, Mayer et al. [32] proposed a generic typology of trust, consisting of three dimensions: ability, benevolence, and integrity. These three dimensions are conceptually distinct since they tap into different elements of cognitive and affective abstraction of trust. Yet collectively, they represent a comprehensive yet parsimonious dimension space for trust formation. Many of the other dimensions proposed in the literature can be reconciled within these three dimensions, as shown in Table 2. This typology of trust is used as the theoretical basis for the proposed trust scale.

Ability

Ability refers to the trustor's perception of trustee's competencies and knowledge salient to the expected behavior [32]. Such perceptions may be based on prior (first-hand or secondhand) experience or institutional endorsements (such as M.D. for doctors, Better Business Bureau membership for firms). In e-commerce contexts, user perceptions of firm's ability are based on two related beliefs: (1) whether the firm is competent (expert or skilled) enough to perform the intended behavior, and (2) whether the firm has access to the knowledge required to perform the behavior appropriately. Perceived lack of these beliefs can undermine perceptions of the trustee's ability. Ability is domain-specific in that trustees highly proficient in one area (and trusted with behaviors in that area) may be viewed as having little aptitude or experience in other areas. For instance, we trust doctors with medical treatment but not with retirement planning! Likewise, firms with demonstrated or perceived domain expertise in a certain area (such as banking) are more likely to be trusted by users in that area than those with less perceived expertise. To overcome this perception, newer online firms attempt to "signal" their ability to users by publishing metrics such as number of accounts, customers, items sold, and other transactional information on their Web sites.

Table 2. Mapping of Trust Dimensions

Study	Mayer et al.'s [32] classification			
	Ability	Integrity	Benevolence	Unclassified
Butler [7]	Competence	Fairness, discreetness, integrity, promise fulfillment, loyalty	Receptivity	Availability, consistency, openness
Cummings and Bromiley [11]	—	Keep commitments, negotiate honestly	Avoid taking excessive advantage	—
Doney and Cannon [14]	—	Credibility	Benevolence	—
Gabarro [17]	Functional competence, interpersonal competence, business sense, judgment	Integrity, discreetness	—	Motives, consistency, openness
Hart et al. [20]	—	Shared values	—	Autonomy, openness
Jarvenpaa et al. [22]	Ability	Integrity	Benevolence	—
Johnson-George and Swap [24]	—	Reliability	—	Emotion
Larzelre and Huston [27]	—	Honesty	Benevolence	—
McKnight et al. [34]	Competence	Honesty	Benevolence	Predictability
Rempel et al. [38]	—	Dependability	Faith	Predictability
Sheppard and Sherman [44]	Competence	Discretion, reliability, integrity	Concern, empathy	Predictability, consistency
Zaheer et al. [49]	—	Reliability, fairness	—	Predictability

Integrity

Domain ability is not adequate for building trust, trustees must also be relied on to fulfill obligations to trustors [38]. Integrity refers to the trustor's perception that the trustee will adhere to a set of principles or rules of exchange acceptable to the trustor during and after the exchange [32]. Perceived integrity instills trustor's confidence in trustee behavior and reduces perceptions of risk. In e-commerce contexts, rules of integrity refer to: (1) conduct of online transactions, (2) customer service policies following a transaction, and (3) firm's use of private user information. However, the specific rules are context-dependent. For instance, online retailing rules include timely shipping of products, timely and accurate billing, and maintaining confidentiality of personal information, whereas online brokerage rules include providing accurate quotes, timely execution of client orders, and timely notification of fulfilled trades. Online firms may build integrity perceptions by explicitly stating their rules of exchange (such as, shipping policies, merchandise return policies, customer data privacy policies) on their Web sites, keeping customers informed of any change in rules, and unflinchingly abiding by those rules. Note that adherence to any set of rules is not adequate; such rules must be perceived by the trustor as being fair and reasonable. Integrity is similar to honesty, fairness, credibility, consistency, predictability, reliability, and dependability dimensions proposed in the literature.

Benevolence

Benevolence is the extent to which a trustee is believed to intend doing good to the trustor, beyond its own profit motive [32]. A benevolent trustee would help the trustor, even when the trustee is not required to be helpful or is not rewarded for being helpful. Benevolence introduces faith and altruism in a relationship, which reduces uncertainty and the inclination to guard against opportunistic behaviors. For example, Progressive.com provides comparative quotes from other insurance carriers for home and auto insurance (even when the other quotes may be better than its own), whereas Charles Schwab provides educational seminars for its clients on investment strategies (even when there is a cost of doing so), and are hence trusted more by customers. However, in many e-commerce contexts, it may be difficult to anticipate user needs for designing benevolent services or it may be overly expensive to deliver such services. In such circumstances, benevolent firms should at the very least: (1) demonstrate receptivity and empathy toward users' concerns and needs, and (2) proactively make good-faith efforts to resolve user concerns.

Nomological Trust Model

Assessing the nomological validity of any scale requires specifying the construct within a nomological network of antecedent and consequent variables, in order to examine the predictive ability of the focal scale. Drawing from prior research, this section proposes such a network by postulating users' familiarity with trustee as a

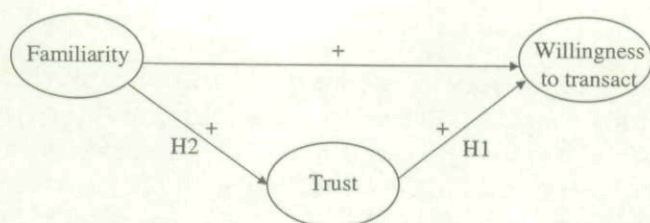


Figure 1. Research Model of Trust

determinant and willingness to transact as a consequence of user trust in online firms (see Figure 1).

Trust is conceptualized in this study in terms of trustor's beliefs in the trustee's ability, benevolence, and integrity. Mayer et al. [32], McKnight et al. [33, 34], and others posit such beliefs to result in corresponding trusting intentions (the intent to engage in trusting behaviors). Conceptual distinction and causality between beliefs and intentions are derived from Fishbein and Ajzen's [15] typology of beliefs, attitude, and intention in the social psychology literature. Although attitude mediate the impact of beliefs on intention in this typology, it is excluded from the proposed nomological network because: (1) the direct effect of beliefs in information systems contexts are generally higher than their indirect effect via attitude [13], and (2) affect embodied in attitude is partially captured within the benevolent and integrity dimensions of trust. In consumer-based e-commerce contexts, trusting intention represents users' willingness to engage in subsequent transactions with online firms [23]. Higher levels of trust in a firm will therefore lead to greater intention on the part of users to engage in online transactions. Hence,

H1. An individual's trust in an online firm is directly related to their willingness to transact with that firm.

It should be noted that trust is only one of several possible reasons for individuals to transact with online firms. One may interact with online firms because it is more convenient than most offline firms, because it offers a broader menu of products and services than offline firms, or because of other reasons. These reasons may be viewed as rival hypotheses to H1 in that a person who does not see adequate value in an online firm or does not perceive it to be useful may not transact with that firm irrespective of their level of trust in that firm. The presence of rival hypotheses ensures the possible falsifiability of H1, a critical requirement for statistical testing of hypotheses.

Familiarity is a widely recognized predictor of trust in the literature [18, 31]. Familiarity refers to one's understanding of another's behavior based on prior interactions or experiences. Although familiarity deals with *past* behaviors, trust refers to expectations about *future* behaviors [18]. Consistency of past behaviors helps trustors esti-

mate the likelihood of desired future behaviors, and hence reduces uncertainty in transactional relationships [50]. The trustor's trust in the trustee continuously evolves with time as the trustor becomes familiar with the trustee's behavior, and hence familiarity fosters a process-based or relational trust among transacting parties. This process model of trust formation, developed via a temporal pattern of cooperative behavior, has been demonstrated analytically in game theoretic models using repeated decision games (for example [12]). Note, however, that familiarity-based trust assumes prior favorable behavior by the trustee; unfavorable or unexpected behaviors may conversely lead to disruption or breakage of trust. Therefore,

H2. An individual's familiarity with an online firm is directly related to their level of trust in that firm.

Note that familiarity is only one of several possible means of building trust. This relational perspective fails to explain *initial* trust or the formation of trust before any transaction has taken place [34]. Under those circumstances, other means of forming trust, resulting from calculative, identification, and institutional beliefs, can become more prevalent [32, 50]. For instance, one's trust in an unfamiliar online firm may be based in their belief that firms generally do not misappropriate consumer trust because of strategic or legal considerations (institutional belief), because the risks of breaking trust outweigh the benefits that firms may incur from that behavior (calculative beliefs), or because firm managers belong to the same race, ethnic background, professional background, and so forth, as the trustor and is therefore not likely to defraud the trustor (identification beliefs). These alternative means of trust formation can be viewed as rival hypotheses for the proposed H2.

Since past behavior is often a good predictor of future behaviors, familiarity can also be expected to have a direct effect on a trustor's willingness to transact with the trustee, apart from its indirect effect via trust [18]. Prior successful transactions with an online firm may motivate users to unthinkingly engage in transactions with that firm, without a conscious cognitive assessment of the trustee's ability, benevolence, and integrity. Hence, for purposes of completeness, the proposed nomological network includes an association between familiarity and willingness to transact (see Figure 1). However, this association is not stated as a formal hypothesis since it does not directly involve the trust construct.

The aforementioned hypotheses are relevant from a practitioner's perspective, since they present trust as a viable means of improving transaction levels or sales volume of online firms. Admittedly, the nomological model discussed above presents a very simplistic view of online relationships and ignores several other facets of such relationships such as cost, breadth of offerings, ease of use, and the like. Other antecedents and consequents of trust proposed in the trust literature are not included in this model, because its purpose is only to test the nomological validity of the proposed trust scale, rather than to develop a complete model of trust. Authors interested in a comprehensive theory of trust should, however, consider including additional variables to build a richer and holistic trust model.

Initial Scale Construction

Item Selection

PSYCHOMETRICIANS EMPHASIZE THAT THE VALIDITY of a measurement scale should start with initial item construction. As Nunnally pointed out, "Rather than test the validity of measures after they have been constructed, one should ensure the [content] validity by the plan and procedures for [instrument] construction" [36, p. 258]. Nunnally suggested a "domain sampling" method, which assumes a predefined domain of content for each measurable construct, and recommends that researchers start with a conceptual definition of the construct of interest, explicate its domain of content, and select candidate items that can faithfully represent this domain. Accordingly, a three-step procedure was employed to create items for the proposed trust scale and establish its content validity. First, salient substrata or facets of each of the three trust dimensions were identified by conceptualizing them against the e-commerce context. Second, items from prior trust literature reflecting each dimension of trust were identified and reconciled to minimize semantic overlap across items. Third, prior scale items that matched best with the desired trust substrata were selected and reworded to relate specifically to the e-commerce context. New items were created to represent substrata that were not represented in prior literature. The combination of top-down substrata specification and bottom-up item matching approaches ensured that proposed trust items adequately represented the entire domain of e-commerce trust, whereas minimizing any extant biases in current trust scales (such as, excess coverage in some areas of the content domain and under-coverage in others).

The optimal number of items for the proposed trust scale was determined using the Spearman-Brown prophecy formula. This formula estimates the number of items needed in a scale to obtain a desired reliability based on the number of items and reliabilities of comparable prior scales. Extrapolating from the prior trust scales (albeit not in e-commerce settings), the formula suggested that six to seven items are required for a scale reliability of 0.80. Reliability of 0.80 is generally considered adequate for confirmatory research (cf. [13]), and some researchers (for example [6]) contend that reliabilities exceeding 0.80 may suggest redundancy in scale items. Further, in applied testing situations where multiple constructs are being measured, it is important to keep measurement scales as small as possible to reduce semantic overlap between scales and minimize respondent fatigue. Allowing for item elimination during subsequent scale validation, it was decided that the initial scale should include three items per trust dimension plus an overall trust item, for a total of ten items.

The first phase of scale construction required specifying three items for each of the three hypothesized trust dimensions: ability, integrity, and benevolence. As discussed earlier, an online firm's ability refers to whether or not users perceive that firm as: (1) possessing necessary domain-specific skills (or expertise), and (2) having access to appropriate information required to perform the expected behavior. Hence, two of the three ability items should refer to skills and information substrata of the ability

dimension, whereas the third item may reflect overall perceptions of a firm's ability. Likewise, users judge a firm's integrity based on the perceived fairness in its rules of exchange in (1) the conduct of online transactions, (2) customer service policies following a transaction, and (3) the use of private user information. Hence, integrity items of the proposed trust scale should reflect these three integrity substrata. Finally, an online firm's benevolence is identified as whether or not the firm (1) demonstrates receptivity and empathy toward users' concerns and needs, and (2) proactively makes good-faith efforts to resolve user concerns. Two of the three benevolence items should reflect these substrata, plus a third item assessing overall firm benevolence (that is, whether it acts in users' best interests).

In the second phase, 14 published trust scales were reviewed [7, 10, 11, 14, 20, 22, 23, 24, 27, 35, 38, 39, 41, 49]. Several of these scales did not address the specific trust dimensions of interest to this study (such as, Sako and Helper's [41] eight-item scale measured trust, distrust, and opportunism; Nooteboom et al.'s [35] six-item scale measured habitualization and institutionalization) or addressed them partially (such as, Cummings and Bromiley's [11] 62-item scale spanned keeping commitments, honesty in negotiations, and avoid taking excessive advantage; Rempel et al.'s [38] 26-item trust scale assessed dependability, predictability, and faith). Some scales measured trust in interpersonal (for example [38]) or interdepartmental (for example [11]) contexts, and only one [23] examined trust in an e-commerce context. Some scales (for example [24]) were not validated, probably because their focus was on hypotheses testing rather than instrument development. Further, several scales used multiple adjectives with similar meaning to explicate the underlying trust dimension (such as, reliability, honesty, and fairness to represent the integrity dimension [11]). Although this approach may increase scale reliability, it may not always capture the entire range of facets (substrata) of trust, resulting in high measurement error. In order to provide representational coverage for the entire content domain, it was decided that one item per substratum would be included in the initial trust scale.

From prior trust scales, items that met three criteria were filtered out: (1) they examined one of the three hypothesized dimensions of trust, (2) they could be adapted to assessing user trust in online firms, and (3) they did not overlap with any of the study's other constructs (that is, willingness to transact and familiarity). Items with substantial semantic overlap were reconciled into a single item. For instance, the items "I trust this store keeps my best interests in mind" [22], "I have faith in my contact person to look out for my best interests even when it is costly to do so" [49], and "We trust this vendor keeps our best interests in mind" [14] were grouped into one single "benevolence" (best interest) item. Negatively worded items in the item pool were reworded positively, since negative items reflect distrust, which is discussed in the literature as an entirely different construct from trust, rather than being a polar opposite [30]. Several popular scales in information systems research are based entirely on positively worded items (such as, Davis's [13] perceived usefulness and ease of use scales, Agarwal and Prasad's [1] personal innovativeness scale), hence, positive valence was not seen as a threat to scale validity. This reconciliation process led to a shortlist of 18 items. Each item was reworded to relate specifically to an

online retailing context (with Amazon.com as the trustee) and anchored using a seven-point Likert scale ranging from "strongly disagree" to "strongly agree."

The final phase in instrument construction involved matching prior scale items to the desired trust substrata and creating new items for missing substrata. A cursory examination of candidate items revealed that they represented only four of the nine hypothesized trust substrata. Substrata with no corresponding items from prior trust scales included access to transaction-related information, fairness in the conduct of online transactions, fairness in using private information, empathy, and resolving user concerns (see Table 3). Although unexpected, this was reasonable since the constituents of user trust in online firms were somewhat different from that in interpersonal, intergroup, or interfirm relationships (the typical contexts for prior trust scales). Single items were added for each of the missing substrata based on their conceptualizations (see Table 3). The initial item pool consisted of 18 items from prior scales plus five items for the missing substrata for online trust.

Pretest and Item Refinement

The next goal was to reduce the initial item pool to nine items representing each of the nine substrata of trust, verify the scale's content validity, and refine individual items if needed. The overall trust item (T10 in Table 3) was dropped from this pool because it did not represent any specific substratum. Item reduction and refinement was done by a pretest sample of six online users, all of whom had purchased at least one product online within the previous year. Each participant was provided with operational definitions of the nine trust substrata and 23 initial trust items printed on separate index cards. Participants were asked to note ambiguity in item wording (and recommend changes if necessary), categorize individual items into nine substrata, and rank items within each substratum according to their semantic proximity to that of the underlying substratum. Similarity across participant responses were assessed via the Q-sort technique, a widely employed technique for confirming the underlying structure of complex variables and establishing their content validity [43].

Based on participant feedback, two items were rephrased for further clarity and three items were eliminated on grounds of having substantive semantic overlap with other items in the candidate item pool. For the remaining items, participants individually classified each item into one of the nine trust substrata. On average, 85 percent of items classifications were correctly classified by the respondent panel. The proportion of correct classification ranged between 71 percent (for the benevolence substratum) to 97 percent (for the expertise substratum) for the multiple-item substrata, and between 83 percent and 100 percent for the single-item substratum (not derived from prior scales). For the overall classification, Cohen's kappa (a measure of inter-rater reliability) was estimated as 0.82. The high rates of correct classifications confirmed that individual items exhibited relatively consistent meaning across the participant panel, and hence multiple Q-sorts or additional statistical analyses were not deemed necessary.

Finally, to select one item to represent each of the four multiple-item substrata, participants were asked to rank order the items within each substratum based on their

Table 3. Initial Items for the Trust Scale

Substratum	Item	Source
Dimension: Ability		
T1. Expertise	Amazon [OBD] has the skills and expertise to perform transactions in an expected manner.	[22]
T2. Information	Amazon [OBD] has access to the information needed to handle transactions appropriately.	New item added
T3. Overall ability*	Amazon [OBD] has the ability to meet most customer needs.	[22]
Dimension: Integrity		
T4. Fairness in transactions	Amazon [OBD] is fair in its conduct of customer transactions.	[11]
T5. Fairness in data usage*	Amazon [OBD] is fair in its use of private user data collected during a transaction.	New item added
T6. Fairness in service	Amazon [OBD] is fair in its customer service policies following a transaction.	New item added
Dimension: Benevolence		
T7. Empathy	Amazon [OBD] is open and receptive to customer needs.	New item added
T8. Benevolence*	Amazon [OBD] keeps its customers' best interest in mind during most transactions.	[14]
T9. Resolving concerns	Amazon [OBD] makes good-faith efforts to address most customer concerns.	New item added
Dimension: Overall trust		
T10. Overall trust	Overall, Amazon [OBD] is trustworthy.	[23]

* Items dropped from final scale.

closeness in meaning with that of the underlying substratum. For this purpose, participants were encouraged to consult with the substratum definitions as frequently as needed. Ranks for each item across the participant panel were summed, and the item with the lowest rank within each substratum was chosen to represent that particular substratum. The selected items are listed in Table 3. Given the small size of the pretest sample, no additional statistical analyses were performed. With the addition of an item to assess overall trust, the pretest led to an initial 10-item trust scale.

Additional scales were required to measure the willingness to transact and familiarity constructs also specified in the nomological model of trust. Both of these scales were adapted directly from prior research. Willingness to transact was measured using three items taken from Jarvenpaa et al.'s [23] willingness to buy scale, whereas the four familiarity items were taken from Gefen [18]. These items are provided in Table 4.

Scale Validation and Refinement

PRELIMINARY PSYCHOMETRIC PROPERTIES of the initial trust scale were examined via a pilot test using MBA students. The results were used to refine and streamline the

Table 4. Items for Other Scales

Construct	Item	Source
Willingness to transact	W1. I intend on using Amazon [OBD] for some of my future purchases [banking transactions].	[23]
	W2. I am inclined to purchase Amazon's [OBD's] goods and services.	
	W3. I am likely to utilize the services provided by Amazon [OBD].	
Familiarity	F1. I am familiar with the process of searching for books [reconciling transactions] on Amazon [OBD].	[18]
	F2. I am familiar with buying books [paying bills online] at Amazon [OBD].	
	F3. I am familiar with inquiring about book ratings [account information] on Amazon [OBD].	
	F4. Overall, I am familiar with Amazon [OBD].	

initial scale, which was then revalidated using data from a field survey of online banking users. Results of these studies are described next.

Pilot Test

The initial trust scale was administered to a sample of 147 evening MBA students enrolled in an e-commerce class at a large U.S. university. Survey items examined subjects' use of Amazon.com, a widely recognized online retailer of books, compact disks, videos, toys, electronics, and other commonly purchased items. All subjects used the Internet frequently in their everyday (work, school, and personal) life, had previously purchased products or services online, and were therefore considered fairly representative of the typical e-commerce user. All subjects were aware of Amazon prior to the survey, and 29 percent had purchased products from Amazon's Web site prior to the survey date.

Following a tour of Amazon's Web site (which pointed out the firm's various trust-building initiatives such as book ratings, privacy statement, and secure connection for billing information), subjects were asked to record their perceptions of Amazon on an online survey form. This form included the items listed in Tables 3 and 4, plus some demographic variables. The online mode of data collection, although relatively new in information systems research, was appropriate since subjects' online behavior was the focus of this study and the Internet is the most effective way to reach a population of online users. Such online surveys are increasingly being used for research purposes (for example [47]), because of their expediency in data collection, ease of data tabulation, and ability to reach a wide population of users. All respondents indicated that they were comfortable with the process of filling out online forms (mean re-

sponse to a single-item question on subjects' degree of comfort in completing an online survey relative to a paper survey was 6.1 on a seven-point Likert scale); hence, the data collection method was not expected to introduce any novelty bias on survey responses.

Given the theory-driven approach to scale development, the confirmatory factor analysis (CFA) approach was employed for scale validation. CFA combines *ex ante* theoretical expectations with empirical data for factor validation, and is therefore a stronger statistical method than alternative approaches such as exploratory factor analysis. CFA was performed on 74 randomly selected observations from the pilot test data (retaining a holdout sample of 73 for further scale validation), with the item correlation matrix as input and maximum likelihood as the model estimation technique. CFA requires a sample size of five to ten times the number of items, hence the current sample size of 74 was appropriate for 10 items. Skewness for scale items ranged between -0.16 and 0.48 and kurtosis ranged between -0.29 and 0.75 (both within the recommended -2 to +2 range), which suggested reasonably good distributional properties for the empirical data. As an initial check of data quality, standardized residuals among individual scale items ranged between -0.049 and 0.087, well below the recommended threshold value of 3.0.

The initial trust scale was modeled as a second-order factor model, with items T1-T3, T4-T6, and T7-T9 as reflective indicators of latent first-order factors (ability, integrity,* and benevolence dimensions of trust, respectively). In turn, the three first-order factors plus item T10 were reflectively related to the second-order trust construct (see Figure 2). Results of the CFA analysis are provided in Figure 2.

The correspondence between theoretical specification and empirical data in a CFA model is assessed using goodness-of-fit metrics such as χ^2 . However, the χ^2 statistic is particularly sensitive to sample sizes (that is, the probability of model rejection increases with increasing sample size, even if the model is minimally false), and hence adjusted χ^2 (χ^2/df ; df = degrees of freedom) is suggested as a better fit metric [4]. It is recommended that this metric not exceed five for models with good fit [3]. For the current CFA model, χ^2/df was 2.33 ($\chi^2 = 65.26$; $\text{df} = 28$), suggesting acceptable model fit. EQS provides additional goodness-of-fit measures such as Normed Fit Index (NFI), Nonnormed Fit Index (NNFI), and Comparative Fit Index (CFI). NFI is sensitive to sample size and may indicate poor fit with small samples even with correct models [4]. However, NNFI (which is robust to sample sizes) and CFI (which is robust to both sample sizes and the distribution of the χ^2 statistic) are considered to be adequate fit metrics, and should exceed 0.90 for good model fit [3]. NFI, NNFI, and CFI were, respectively, estimated as 0.92, 0.93, and 0.95 for the second-order trust model, and hence, the theorized model fit well with the observed data.

The next step in CFA is to examine hypothesized path coefficients and their significance. It is recommended that measurement factor loadings exceed 0.70, so that each item is explained more by its hypothesized reflective construct than by error [16]. This criterion was met by eight out of ten trust items (see Figure 2), with the exception of items T5 and T8 (0.65 and 0.62 loadings, respectively). An examination of the correlation matrix (see the Appendix) revealed that the overall trust item (T10) had

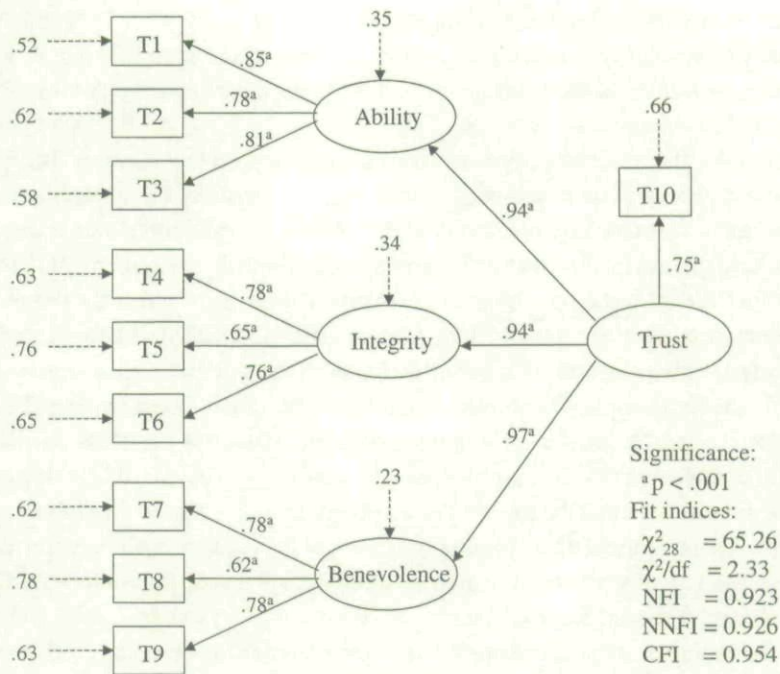


Figure 2. Second-Order CFA Model with Ten Trust Items

relatively poor correlations with items T5 and T8 (in the 0.4–0.5 range) compared to those with other items (greater than 0.5), suggesting scale unidimensionality problems for these two items. Item T5 examined subjects' perception of how fair Amazon was in its usage of consumer data, whereas item T8 asked whether Amazon kept consumers' best interests in mind while conducting transactions. Although subjects had some prior cognitive or affective basis to form an informed judgment for other trust items (such as, due to prior experience or otherwise), they probably did not have any means to evaluate Amazon's fairness or concern for customer's interests, and hence may have underweighted these items. It is also possible that for well-known firms such as Amazon, these issues may be viewed by consumers as being redundant or subsumed within other scale items. Hence, it was decided to drop items T5 and T8 from the trust scale. Dropping these items enhanced the average inter-item correlation from 0.54 to 0.57, and dropped the Cronbach alpha for the overall trust scale minimally from 0.93 to 0.92.

However, dropping T5 and T8 from the trust scale decreased the proportionate weighting of the integrity and benevolence dimensions (relative to the ability dimension) on the overall trust scale. In order to relatively equal weighting for each of the three trust dimensions, it is decided to also drop one ability item. Of the three ability items, items T1 and T2 represented two aspects of the trustee's ability (skills and access to information respectively), but item T3 did not connote any specific aspect of ability (it assessed overall ability, which partially overlapped with T1 and T2).

Given the redundancy implicit in T3, this item was selected for dropping from the overall trust scale.

It may be interesting to note that none of the remaining seven trust items are highly specific to Amazon. This is partially by design and partially due to the nature of the trust construct. First, the original items were worded in a way to ensure that the scale can be equally applied to a wide range of online contexts (online retailing or otherwise). In fact, the current study intended to create a trust scale that can be applied equally to online retailing and online banking. Second, given the anonymity involved in online firms, it is possible that consumers' trust perceptions of online firms are based on similar facets, irrespective of the industry of that firm. However, each of the seven remaining trust items represented a belief that is critical to trust formation (such as, firm's skills and expertise, access to information, fairness in transactions, openness and receptivity, and so on) yet are not directly observable by virtue of the online nature of these firms.

A more critical observation was the unusually high path coefficients between trust and its three dimensions (ability, integrity, and benevolence) in the second-order factor model. The standardized path coefficient for these paths ranged between 0.94 and 0.97. Such high path coefficients are indicative of multicollinearity, implying that the three trust dimensions are empirically inseparable even though they may be conceptually distinct. In other words, a user who rates an online firm highly on its ability is also likely to perceive that firm as being high in integrity, benevolence, and overall trust. It is possible that a subjects' overwhelming familiarity with Amazon may have led them to believe that the online firm is excellent in all three dimensions of trust and that an examination of a less familiar firm may have teased out the effects of the three dimensions better. At any rate, multicollinearity suggested that it is more appropriate to model trust as a first-order construct (reflectively linked to items from the three dimensions), rather than as second-order construct.

Consequently, the pilot test data was reexamined using the holdout sample of 73 Amazon users, after dropping items T3, T5, and T8 and modeling trust as a first-order construct. Results of this analysis are presented in Figure 3. Indeed, this model demonstrated a slightly better fit ($NFI = 0.94$, $NNFI = 0.95$, $CFI = 0.96$) compared to the second-order model. All of the seven paths in the measurement model were significant and had loadings exceeding 0.70. Good model fit, combined with high factor loadings, provided initial evidence of scale unidimensionality. Although both the first- and second-order representation of trust fit well with observed data, the first-order model is preferred by virtue of its simplicity and slightly better fit, and is therefore retained for further analysis.

Confirmatory Study

Although the pilot test validated the items and factor structure of the trust scale, it provided little evidence of convergent and discriminant validity and predictive ability (nomological validity) of the proposed scale. Moreover, it is unknown whether the proposed scale is generalizable to other online (non-retailing, nonstudents) settings.

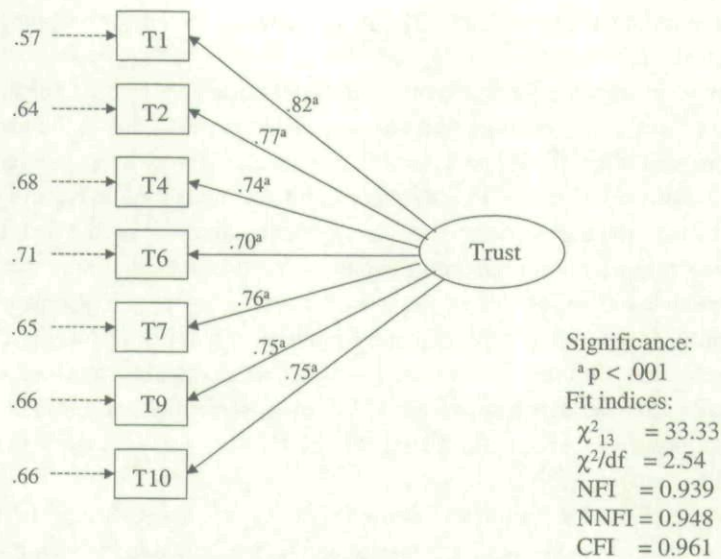


Figure 3. First-Order CFA Model with Seven Trust Items

A second confirmatory survey of online banking users was conducted to examine these issues.

Subjects for the second study were customers of the online banking division (OBD) of one of the largest national banks in the United States. OBD offered its customers an integrated suite of personal banking services (online checking and savings accounts, money market accounts, certificate of deposits, credit cards, home equity loans, home mortgage, insurance, investment services, portfolio management, and retirement planning services), higher yields compared to traditional offline accounts, the flexibility and convenience of 24-hour banking, and an array of sophisticated financial tools (such as, online bill payment, online funds transfer, transaction register, customized reporting). The choice of a single data collection site controlled for the potential effects of macro-level firm variables (such as, customer interface, infrastructure) on user behavior, thereby increasing the likelihood of detecting micro-level individual effects better.

The subject sample consisted of 1,000 randomly selected online customers from OBD's existing customer base of over 1 million users. A message soliciting their participation in a survey of online banking practices was electronically mailed to these subjects by OBD, appended to one of their regular monthly statements (also distributed via e-mail). The message outlined the purpose of the study, provided a hyperlink to an online survey form (consisting of trust, familiarity, and willingness to transact items), and as an incentive, offered respondents the opportunity to register in a drawing of small cash prizes. Customer responses were edited by OBD to remove any identifying data (such as, e-mail addresses, IP addresses, and user name) before making them available to the researcher. A total of 122 usable responses were returned for a response rate of 12 percent. The low response rate may be attributed to

the fact that the solicitation message was attached to the end of a subjects' monthly statements, and therefore, may have been overlooked by some subjects. Further, no follow-up request was sent to nonrespondents, since OBD wanted to avoid the risk of being perceived as a "spammer."

The respondent group for the second study was 38 percent female, had a mean age of 33.7 years, had a mean annual income of \$75,000, and subscribed to a wide range of professions (students, professionals, self-employed, academics, executives, retirees). Subjects maintained an average online balance of \$7,000, 92 percent of subjects also maintained a traditional bank account in addition to their online account, and had maintained their online account for an average of eight months (at the time of the survey). Compared to the Amazon study, OBD users exhibited greater levels of trust and greater willingness to transact. Item means and standard deviations for this sample are listed in Table 5.

As a check of data quality, standardized residuals for the confirmatory study ranged between -0.102 and 0.114, again within recommended limits. The observed data met tests of univariate normality, which suggests (though not guarantees) multivariate normality. Further, skewness and kurtosis estimates ranged well within the -2 to +2 range, suggesting no serious departures from the normality assumption.

Convergent and discriminant validity and scale unidimensionality were assessed via the CFA approach (performed using EQS). However, in contrast to the pilot test, all three scales in the proposed nomological trust model (trust, familiarity, and willingness to transact) were included in this analysis. Each item was modeled as a reflective indicator of its hypothesized latent construct, and the three constructs were allowed to covary. The trust scale consisted of seven items and was modeled as a first-order construct. This CFA model demonstrated excellent model fit. Model χ^2 was 114.07 ($df = 74$) for a χ^2/df of 1.54, and NFI, NNFI, and CFI were 0.89, 0.95, and 0.96, respectively (see Table 5). Inter-factor correlations were between 0.44 and 0.58 (see Table 6).

Convergent validity was evaluated for the three constructs using three criteria recommended by Fornell and Larcker [16]: (1) all measurement factor loadings must be significant and exceed 0.70, (2) construct reliabilities must exceed 0.80, and (3) average variance extracted (AVE) by each construct must exceed the variance due to measurement error for that construct (that is, AVE should exceed 0.50). As evident from the measurement model in Table 7, all indicator factor loadings (λ) were significant at $p = 0.001$ (lowest t -statistic being 9.3) and the lowest λ value was 0.74. Composite reliabilities (ρ_c) of constructs ranged between 0.83 and 0.91, and Cronbach alphas ranged between 0.83 and 0.89 (see Table 6). AVE ranged from 0.59 to 0.64 for the three constructs (see Table 6). Hence, all three scales met the norms for convergent validity.

Discriminant validity was assessed using a series of χ^2 difference tests, where the χ^2 statistic of the unconstrained CFA model (where all constructs are freely correlated) is compared with that of constrained models where the correlation between pairs of constructs is set to one. Significant differences in χ^2 between the unconstrained and constrained models provide proof of discriminant validity between the constrained

Table 5. CFA Measurement Model Results

Item	Mean	Standard deviation	Standardized item loading (λ) ^a	Standardized error loading (ϵ)	t-Statistic (for λ)
W1	4.89	1.02	0.811	0.585	9.962
W2	5.21	1.10	0.732	0.681	8.712
W3	4.96	0.97	0.814	0.581	10.009
T1	4.56	1.12	0.810	0.587	10.525
T3	5.02	1.15	0.760	0.650	9.580
T4	4.67	1.10	0.744	0.668	9.302
T6	4.93	1.07	0.762	0.648	9.622
T7	5.11	1.15	0.753	0.658	9.465
T9	4.78	1.10	0.783	0.623	10.002
T10	5.05	1.10	0.775	0.632	9.860
F1	4.31	1.21	0.812	0.584	10.383
F2	4.39	1.15	0.817	0.576	10.484
F3	4.64	1.01	0.832	0.555	10.760
F4	4.12	0.97	0.759	0.651	9.423

Notes: W = willingness to transact, T = trust, F = familiarity; Model fit: $\chi^2_{74} = 114.07$, $\chi^2/df = 1.541$, NFI = 0.890, NNFI = 0.948, CFI = 0.958. ^a All item loadings (λ) were significant at $p = 0.001$ level.

Table 6. Scale Properties and Factor Correlations

Construct	Number of items	Cronbach alpha	Composite reliability ^a	AVE ^b	Factor correlations		
					W	T	F
W	3	0.83	0.83	0.62	1.000	—	—
T	7	0.89	0.91	0.59	0.542	1.000	—
F	4	0.88	0.88	0.65	0.577	0.441	1.000

Notes: W = willingness to transact, T = trust, F = familiarity. ^a Composite reliability $\rho_c = (\sum \lambda)^2 / [(\sum \lambda)^2 + \sum \text{var}(\epsilon)]$. ^b AVE = average variance extracted.

Table 7. Chi-Square Tests of Discriminant Validity

Variables constrained	Chi-square	Degrees of freedom	Chi-square difference ^a
None	114.07	74	—
W, T	206.25	75	92.18
W, F	197.33	75	83.26
T, F	311.54	75	197.47

Notes: W = willingness to transact, T = trust, F = familiarity. ^a All χ^2 differences were significant at $p = 0.001$ level.

pair of constructs [2]. As shown in Table 7, χ^2 differences ranged between 92.18 and 197.47. All χ^2 differences were significant at $p = 0.001$, demonstrating adequate discriminant validity for the three hypothesized scales. As a stronger test of discriminant validity, Fornell and Larcker [16] suggested that the average variance extracted for each construct should exceed the squared correlation between that and any other constructs. As seen from the factor correlation matrix in Table 6, the lowest AVE was 0.59 (for trust scale), which exceeded the largest squared correlation between any pair of constructs (0.33 between willingness to transact and familiarity). Hence, the latter test of discriminant validity was also met.

Finally, nomological validity of the trust scale was assessed in this study by testing hypotheses H1 and H2 specified earlier within the simple trust model presented in Figure 1. These hypotheses were collectively tested using EQS structural equation modeling. For this purpose, the covariance between familiarity, trust, and willingness to transact in the previous CFA model were replaced by causal associations depicted in Figure 1. Note that the link between familiarity and willingness to transact was also included for purpose of model completeness, although this link was not of focal concern to this study. The goodness-of-fit metrics, path coefficients, and path significance for this structural model are presented in Figure 4.

The hypothesized research model exhibited good fit with observed data ($\chi^2/df = 1.61$, NFI = 0.89, NNFI = 0.94, CFI = 0.96), comparable to that of the previous CFA. Item loadings in the measurement model were also similar to those in the CFA. Of greater interest for nomological validity were the path estimates in the structural model and variance explained (R^2 value) in each dependent variable. All three hypothesized paths were supported at the 0.001 significance level (see Figure 4), providing support for hypotheses H1 and H2. Both trust and familiarity were significant predictors of users' willingness to transact ($\beta = 0.36$ and 0.42 , respectively), which collectively explained 31 percent of the variance in the dependent variable. Familiarity was also a significant predictor of trust ($\beta = 0.44$), and explained 19 percent of the trust variance. These effects, although not substantially large, demonstrates that the final seven-item trust scale does indeed measure what it is purported to measure (that is, users' willingness to transact with an online firm) and is predicted by theorized determi-

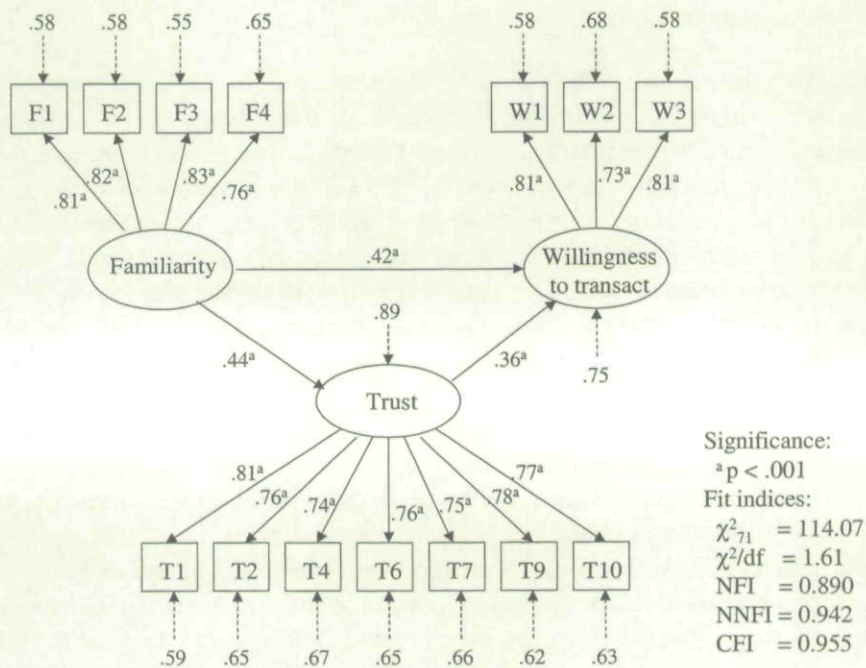


Figure 4. Structural Equation Model Analysis of Online Banking

nants (that is, familiarity), thereby satisfying the nomological validity requirement of the proposed trust scale.

However, the large proportion of variance left unexplained in each dependent variable suggests that there may be other predictors that are missing from the current nomological model. For instance, one's willingness to transact with an online firm may be predicted by additional variables above and beyond trust, such as a perceived usefulness and perceived ease of use of such transactions [13]. Likewise, one's trust in an online firm may not necessarily be derived only from prior familiarity with the firm, but also from calculative, institutional, and identification beliefs about the firm [32]. The latter argument is particularly true for new firms, with no prior history of transaction or familiarity. At the same time, the low associations between this study's constructs may also be an artifact of the timing and context of the online relationships examined in this study (Amazon and OBD). In both cases, users had some prior familiarity or transactional history with the online firm, which may have led to a high level of trust preceding this study. Lack of variability in the independent variables (familiarity and trust) may have resulted in masking the true effects of these variables on the dependent variables (trust and transaction intention). In order to establish the true nature of these relationships, one has to devise a research design incorporating familiar and unfamiliar firms (or examine trust longitudinally in one online firm as subjects gain familiarity with its usage) to ensure adequate variability in the constructs and detect the hypothesized effects better.

Conclusions and Limitations

THE PURPOSE OF THIS PAPER WAS TO DEVELOP and validate an instrument for measuring individual trust in online firms, for potential use in future empirical studies of trust in e-commerce contexts. Scale construction is one of the most important steps in confirmatory research because the quality of a measurement scale determines the extent to which empirical results are meaningful and accurate. As discussed earlier, prior trust scales emphasized contradictory conceptualizations of trust and were not directly applicable to e-commerce settings. Based on three core dimensions of trust (ability, integrity, and benevolence) their underlying substrata in e-commerce contexts, an initial 10-item domain-specific trust scale was constructed. Multiple rounds of empirical validation via field surveys of online retailing and online banking users helped refine this scale into a seven-item scale that demonstrated adequate psychometric properties and predictive ability.

However, scale development is hardly a research objective in its own right, if the designed scale is not employed in subsequent studies. As such, the development of a trust scale in this paper was part of a larger study examining the dynamics of trust formation in online relationships (to be reported elsewhere). Future studies investigating consumers' Web shopping, purchase, or interaction behaviors may use the proposed scale to evaluate the role played by trust in such contexts. The proposed scale may be useful in examining consumer trust in online entities (that is, firms) versus the online media (that is, Internet). Finally, one of the most interesting and value-added areas of online commerce is business-to-business relationships, where trust is a substantive issue [49]. Buying firms must trust selling firms in online exchanges, procurement personnel must trust supplying vendor personnel in online procurement, retailers must trust suppliers replenishing their store shelves if they wish to implement vendor-managed inventory, and product designers and manufacturers from multiple firms must trust each other in online collaborative supply chains. Although the proposed scale was designed in a consumer-based e-commerce context, it may be extended with minor modifications to business-to-business e-commerce contexts characterized by anonymous relationships, lack of control, and potential opportunistic behaviors, which is primarily based on relationship or familiarity. However, note that alternative modes of trust creation in business-to-business relationships (such as, using a letter of credit to form institutional trust) may sometimes limit the applicability of the proposed scale.

From a practitioner standpoint, the trust scale presented here provides a convenient means for e-commerce managers to assess their customers' level of trust in their firm. As demonstrated in this study, trust is a significant predictor of users' willingness to transact with online firms. Firms whose revenue structure depends on frequent and continuous user transactions (such as, online banking, online brokerage) stand to lose most if they fail to assess consumer trust in their services or engage in proactive trust-building initiatives. Further, users with low trust levels are likely to be future discontinuers of the service, and hence, early identification of these users may be the first step to ensuring their retention by targeting them with special marketing or education programs to help rebuild their trust before it is too late.

However, this study is not without limitations. First, one may question whether the entire domain of trust dimensions was represented in the proposed scale. Content validity was assured in this study via an extensive literature review to uncover all possible trust dimensions across different situational contexts (see Table 1), careful reconciliation of these dimensions (see Table 2), and examining the relevance of each dimension within e-commerce contexts. The resulting scale is not only theoretically justified, but also captures salient trust dimensions across a wide range of contexts. Content validity was verified via a pretest using participant panel and Q-sorting. Second, given that most current surveys are still paper or telephone-based, it is possible that the novelty of online surveys may have biased a subjects' response to the survey. To test for such "novelty effect," a single-item measure assessing subjects' degree of comfort with filling out online forms was included in each of the two surveys. As stated before, the mean score on this item was 6.1 on a seven-point scale, suggesting that the novelty of online surveys did not bias subjects' responses in this study. Third, this study found that trust explains only 13 percent of user's willingness to transact and 19 percent of trust is explained by familiarity with the trustee. The large proportion of unexplained variance in these variables suggests the presence of additional determinants that were beyond the scope of this study. However, identifying such determinants and examining their relative effects on the dependent variable represent potentially fruitful ways of advancing the current state of trust research in e-commerce.

REFERENCES

1. Agarwal, R., and J. Prasad. A conceptual and operational definition of personal innovativeness in the domain of information technology. *Information Systems Research*, 9, 2 (1998), 204-215.
2. Bagozzi, R.P.; Yi, Y.; and Phillips, L.W. Assessing construct validity in organizational research. *Administrative Science Quarterly*, 36, 3 (1991), 421-458.
3. Bentler, P.M. *EQS Structural Equations Program Manual*. BMDP Statistical Software, Los Angeles, 1989.
4. Bentler, P.M., and Bonnett, D.G. Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 3 (1980), 588-606.
5. Boudreau, M.C.; Gefen, D.; and Straub, D.W. Validation in information systems research: A state-of-the-art assessment. *MIS Quarterly*, 25, 1 (March 2001), 1-16.
6. Boyle, G.J.; Stankov, L.; and Cattell, R.B. Measurement and statistical models in the study of personality and intelligence. In D.H. Saklofske and M. Zeidner (eds.), *International Handbook of Personality and Intelligence*. New York: Plenum Press, 1995, pp. 417-446.
7. Butler, J.K. Jr. Toward understanding and measuring conditions of trust: Evolution of a conditions of trust inventory. *Journal of Management*, 17, 3 (1991), 643-663.
8. Cheskin Research. eCommerce trust study. Research report, Redwood Shores, CA, January 1999.
9. Compeau, D.R., and Higgins, C.A. Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, 19, 2 (June 1995), 189-211.
10. Couch, L.L.; Adams, J.M.; and Jones, W.H. The assessment of trust orientation. *Journal of Personality Assessment*, 67, 2 (1996), 305-323.
11. Cummings, L.L., and Bromiley, P. The organizational trust inventory (OTI): Development and validation. In R.M. Kramer and T.R. Tyler (eds.), *Trust in Organizations*. Thousand Oaks, CA: Sage, 1996, pp. 302-330.

12. Dasgupta, P. Trust as a commodity. In D.G. Gambetta (ed.), *Trust*. New York: Basil Blackwell, 1988, pp. 49-72.
13. Davis, F.D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13, 3 (1989), 319-340.
14. Doney, P.M., and Cannon, J.P. An examination of the nature of trust in buyer-supplier relationships. *Journal of Marketing*, 61, 2 (1997), 35-51.
15. Fishbein, M., and Ajzen, I. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Boston: Addison Wesley, 1975.
16. Fornell, C., and Larcker, D.F. Evaluating structural equations with unobservable variables and measurement error. *Journal of Marketing Research*, 18, 2 (1981), 39-50.
17. Gabarro, J.J. The development of trust influence and expectations. In A.G. Athos and J.J. Gabarro (eds.), *Interpersonal Behavior: Communication and Understanding in Relationships*. Upper Saddle River, NJ: Prentice Hall, 1978, pp. 290-303.
18. Gefen, D. E-commerce: The role of familiarity and trust. *Omega*, 28, 6 (2000), 725-737.
19. GVU Center. GVU's eighth WWW user survey. GVU Center, Atlanta, 1998. Available at www.gvu.gatech.edu/user_surveys/survey-1997-10.
20. Hart, K.M.; Capps, H.R.; Cangemi, J.P.; and Caillouet, L.M. Exploring organizational trust and its multiple dimensions: A case study of General Motors. *Organizational Development Journal*, 4, 2 (1986), 31-39.
21. Hoffman, D.L.; Novak, T.P.; and Peralta, M. Building consumer trust online. *Communications of the ACM*, 42, 4 (1999), 80-85.
22. Jarvenpaa, S.L.; Knoll, K.; and Leidner, D.E. Is anybody out there? Antecedents of trust in global virtual teams. *Journal of Management Information Systems*, 14, 4 (Spring 1998), 26-64.
23. Jarvenpaa, S.L.; Tractinsky, N.; and Vitale, M. Consumer trust in an Internet store. *Information Technology and Management*, 1, 1-2 (2000), 45-71.
24. Johnson-George, C., and Swap, W.C. Measurement of specific interpersonal trust: Construction and validation of a scale to assess trust in a specific other. *Journal of Personality and Social Psychology*, 43, 6 (1982), 1306-1317.
25. Keen, P.G.W. Are you ready for the "trust" economy. *Computerworld*, 31, 16 (April 21, 1997), 80.
26. Kollock, P. The production of trust in online markets. In E.J. Lawler (ed.), *Advances in Group Processes*, vol. 16. Greenwich, CT: JAI Press, 1999, pp. 99-123.
27. Larzelre, R.E., and Huston, T.L. The dyadic trust scale: Toward understanding interpersonal trust in close relationships. *Journal of Marriage and the Family*, 42, 8 (1980), 595-604.
28. Lee, M.K.O., and Turban, E. A trust model for consumer Internet shopping. *International Journal of Electronic Commerce*, 6, 1 (2001), 75-92.
29. Lewicki, R.J., and Bunker, B.B. Developing and maintaining trust in work relationships. In R.M. Kramer and T.R. Tyler (eds.), *Trust in Organizations*. Thousand Oaks, CA: Sage, 1996, pp. 114-139.
30. Lewicki, R. J.; McAllister, D.J.; and Bies, R.J. Trust and distrust: New relationships and realities. *Academy of Management Review*, 23, 3 (1998), 438-458.
31. Luhmann, N. Familiarity, confidence, trust: Problems and alternatives. In D.G. Gambetta (ed.), *Trust*. New York: Basil Blackwell, 1988, pp. 94-107.
32. Mayer, R.C.; Davis, J.H.; and Schoorman, F.D. An integrative model of organizational trust. *Academy of Management Review*, 20, 3 (1995), 709-734.
33. McKnight, D.H., and Chervany, N.L. What trust means in e-commerce customer relationships: An interdisciplinary conceptual typology. *International Journal of Electronic Commerce*, 6, 2 (2001), 35-39.
34. McKnight, D.H.; Cummings, L.L.; and Chervany, N.L. Initial trust formation in new organizational relationships. *Academy of Management Review*, 23, 3 (1998), 473-490.
35. Nooteboom, B.; Berger, H.; and Noorderhaven, N.G. Effects of trust and governance on relational risk. *Academy of Management Journal*, 40, 2 (1997), 308-338.
36. Nunnally, J. *Psychometric Theory*. New York: McGraw-Hill, 1978.
37. Parker-Pope, T. How to tell whether that online drugstore is really a good deal. *Wall Street Journal*, February 16, 2001, B1.

38. Rempel, J.K.; Holmes, J.G.; and Zanna, M.P. Trust in close relationships? *Journal of Personality and Social Psychology*, 49, 1 (1985), 95-112.
39. Rotter, J.B. Interpersonal trust, trustworthiness, and gullibility. *American Psychologist*, 35, 1 (1980), 1-7.
40. Rousseau, D.M.; Sitkin, R.S.; Burt, R.S.; and Camerer, C. Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, 23, 3 (1998), 393-404.
41. Sako, M., and Helper, S. Determinants of trust on supplier relations: Evidence from the automotive industry in Japan and the United States. *Journal of Economic Behavior & Organization*, 34, 3 (1998), 387-417.
42. Segars, A.H. Assessing the unidimensionality of measurement: A paradigm and illustration within the context of information systems research. *Omega*, 25, 1 (1997), 107-121.
43. Segars, A.H., and Grover, V. Strategic information systems planning success: An investigation of the construct and its measurement. *MIS Quarterly*, 22, 2 (June 1998), 139-163.
44. Sheppard, B.H., and Sherman, D.M. The grammars of trust: A model and general implications. *Academy of Management Review*, 23, 3 (1998), 422-437.
45. Shneiderman, B. Universal usability. *Communications of the ACM*, 43, 5 (2000), 85-91.
46. Straub, D.W. Validating instruments in MIS research. *MIS Quarterly*, 13, 2 (1989), 147-169.
47. Tan, M., and Teo, T.S.H. Factors influencing the adoption of Internet banking. *Journal of AIS*, 1, 5 (2000), 1-42. Available at jais.isworld.org/articles/1-5.
48. Webster, J., and Martocchio, J.J. Microcomputer playfulness: Development of a measure with workplace implications. *MIS Quarterly*, 16, 2 (1992), 201-226.
49. Zaheer, A.; McEvily, B.; and Perrone, V. Does trust matter? Exploring the effects of interorganizational and interpersonal trust on performance. *Organization Science*, 9, 2 (1998), 141-159.
50. Zucker, L.G. Production of trust: Institutional sources of economic structure, 1840-1920. In L.L. Cummings and B.M. Staw (eds.), *Research in Organizational Behavior*, vol. 8. Greenwich, CT: JAI Press, 1986, pp. 53-111.

Appendix

Correlation Matrix for Pilot Test (N = 147)

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
T1	1.00									
T2	0.66	1.00								
T3	0.71	0.62	1.00							
T4	0.57	0.58	0.54	1.00						
T5	0.43	0.51	0.47	0.46	1.00					
T6	0.52	0.47	0.51	0.64	0.45	1.00				
T7	0.62	0.61	0.59	0.53	0.51	0.54	1.00			
T8	0.39	0.42	0.45	0.45	0.63	0.47	0.48	1.00		
T9	0.64	0.51	0.58	0.56	0.49	0.52	0.61	0.49	1.00	
T10	0.63	0.61	0.57	0.51	0.46	0.57	0.52	0.44	0.55	1.00

Correlation Matrix for Confirmatory Study (N = 122)

	W1	W2	W3	T1	T2	T4	T6	T7	T9	T10	F1	F2	F3	F4
W1	1.00													
W2	0.58	1.00												
W3	0.66	0.61	1.00											
T1	0.36	0.29	0.32	1.00										
T2	0.40	0.38	0.35	0.69	1.00									
T4	0.32	0.33	0.37	0.63	0.55	1.00								
T6	0.28	0.24	0.31	0.61	0.54	0.63	1.00							
T7	0.31	0.26	0.29	0.57	0.49	0.61	0.62	1.00						
T9	0.40	0.35	0.32	0.63	0.59	0.48	0.57	0.65	1.00					
T10	0.44	0.31	0.27	0.59	0.61	0.55	0.58	0.57	0.67	1.00				
F1	0.37	0.41	0.38	0.32	0.26	0.28	0.31	0.31	0.25	0.32	1.00			
F2	0.44	0.31	0.29	0.24	0.29	0.29	0.26	0.25	0.26	0.27	0.69	1.00		
F3	0.36	0.33	0.38	0.29	0.26	0.27	0.29	0.26	0.29	0.26	0.63	0.71	1.00	
F4	0.38	0.30	0.47	0.25	0.29	0.29	0.24	0.25	0.24	0.30	0.64	0.55	0.66	1.00

Copyright of *Journal of Management Information Systems* is the property of M.E. Sharpe Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.