Perceived trustworthiness of online shops

Oliver B. Büttner1* and Anja S. Göritz2

1Institute of Marketing and Retailing, Georg-August-Universität Göttingen, Germany
2Department of Organizational and Social Psychology, University of Erlangen-Nürnberg, Germany

- Trust has been found to be crucial for consumer behaviour towards online shops. However, existing studies on the role of trust in electronic commerce are mainly based on ad hoc scales to measure trustworthiness, merely rely on self-reported consequences of trust such as intention to buy, and focus on low-risk products.
- In a web-based study, 634 participants interacted with a provider of medical goods, that is a simulated online pharmacy. The study develops and employs a psychometrically sound scale for assessing perceived trustworthiness of online shops. Moreover, it examines the impact of trustworthiness on both consumers’ intended and actual behaviour towards online shops.
- Results show that trustworthiness promotes both intention to buy and actual financial risk taking. Perceived risk was not found to moderate the relationship between trustworthiness and intention to buy. Instead, trustworthiness partially mediated the influence of perceived risk on intention to buy. The results from the scale development challenge multidimensional conceptualizations of trust; comparing this finding with other studies suggests that the duration of the relationship might moderate the dimensionality of trust.

Copyright © 2008 John Wiley & Sons, Ltd.

In electronic commerce, trust is deemed crucial for turning site visitors into buyers. When engaging in online transactions, customers have to rely on the promises given by the online retailer. For example, customers do not know in advance whether an ordered item will turn out exactly as what they wanted (Garbarino and Strahilevitz, 2004; Nitse et al., 2004). Also, customers have to trust that their personal information will not be passed on to other parties, or that their credit card information is safe from hackers (Jarvenpaa et al., 2000; Belanger et al., 2002; Garbarino and Strahilevitz, 2004). As the online environment features many possibilities for fraud (Grazioli and Wang, 2001), customers are generally eager to find out before any purchase whether a particular online shop is trustworthy or not.

Consequently, trust has received a lot of research attention and has been identified as a key driver for the success of electronic commerce (Jarvenpaa et al., 2000; McKnight...
et al., 2002b; Konradt et al., 2003; Stewart, 2003; Trifts and Häubl, 2003; Harris and Goode, 2004; Bart et al., 2005; Schlosser et al., 2006). However, studies have hitherto predominantly focused on transactions under low risk such as the purchases of laptops, books or travels. Trustworthiness, however, is a more salient issue with high-risk transactions, for example if consumers’ health or even their life are at stake as with health-related online services such as online pharmacies or medical advice (Eysenbach et al., 2002; Arrun˜ada, 2004). Due to their lack of medical expertise and the information asymmetry between customer and provider, consumers are unable to assess the quality of services or products before – and in many cases not even after – receiving them (Arrun˜ada, 2004). Furthermore, interacting with a health-related provider regularly involves the production of sensitive information about one’s health status.

In the study at hand, our primary goal was to establish a scale for measuring the perceived trustworthiness of online shops. Previous scales used for assessing trust have mainly been developed ad hoc. Furthermore, most existing studies are merely based on self-reports of participants such as their intention to buy or attitudes towards the store. Therefore, we developed a scale to measure the perceived trustworthiness of an online vendor and validated this scale by both intended and actual behaviour. Furthermore, the study addresses a gap in research on online as well as on offline trust. Although trust is theorized to be tied to perceived risk (Mayer et al., 1995; Delgado-Ballester and Munuera-Alemán, 2001), the relationship between the two constructs has rarely been empirically examined. In this study, we elaborate and examine the links between perceived risk, trustworthiness and intention to buy. Because of the growing use of health-related online services by consumers and the salience of trust in this domain (Fox, 2005; Sillence et al., 2006), we chose the context of an online pharmacy as an example for electronic commerce. In contrast to the majority of the existing studies, which are based on student samples, we drew onto a heterogeneous sample to extend the generalizability of findings in online trust research (Peterson, 2001; Birnbaum, 2004).

**Background**

**Trust and perceived trustworthiness**

As trust is studied in different disciplines such as psychology, sociology and marketing, and in different contexts like organizations (e.g. Mayer et al., 1995), romantic relationships (e.g. Rempel et al., 1985) or buyer–seller relationships (e.g. Ganesan and Hess, 1997), a multitude of definitions of trust exist. However, most scholars agree that trust has three major constituents: (a) uncertainty about the outcomes of an interaction, (b) personal harm as a possible outcome of the interaction and (c) lack of influence on the outcomes (Petermann, 1996). Trust has been conceptualized both as a trait and as a state. A prominent example for trust as a trait is Rotter’s (1967) interpersonal trust, which describes a generalized expectancy about the behaviour of others. Empirical results, however, suggest that in a particular situation the influence of trust as a trait is limited compared to the attributes of the trustee (e.g. Schlenker et al., 1973). Attributes of the trustee are emphasized in state-oriented approaches; they focus on trust towards a specific interaction partner (Kee and Knox, 1970; Morgan and Hunt, 1994; Mayer et al., 1995; McKnight et al., 1998).

Within state conceptualizations of trust, two different notions have evolved (Moorman et al., 1992; Mayer et al., 1995; Gefen et al., 2003; Schlosser et al., 2006). First, trust has been understood as a willingness to take risks in a relationship (Deutsch, 1960; Schlenker et al., 1973; Moorman et al., 1992; Mayer et al., 1995). This is addressed by Mayer et al. (1995, p. 712) when they define trust as ‘the willingness to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party’.
Second, trust has been treated as a trustor’s set of beliefs about certain qualities the trustee possesses or lacks (Kee and Knox, 1970; Mayer et al., 1995; McKnight et al., 1998). These beliefs are cognitive evaluations that result from attribution processes (Fishbein and Ajzen, 1975). Using Mayer et al.’s (1995) terminology, we will refer to this belief-based conceptualization of trust as perceived trustworthiness.

Mayer et al. (1995) provide a framework for integrating these two conceptualizations of trust towards a particular interaction partner. In their model, perceived trustworthiness encompasses three dimensions: ability, benevolence and integrity. Ability refers to the trustee’s competence to fulfil promises given. Benevolence denotes that the trustee is interested in the trustor’s well-being. Integrity means that the trustee follows a set of desirable principles. McKnight et al. (1998) have built upon this model and included a fourth dimension: the predictability of the trustee’s behaviour. Beliefs about these four dimensions of perceived trustworthiness affect the degree of trust (as willingness) towards the trustee (Mayer et al., 1995; McKnight et al., 1998; Mayer and Davis, 1999). Other authors have carved out these dimensions too, but not necessarily under the same names or in the same composition (e.g. Kee and Knox, 1970; Larzalere and Huston, 1980; Doney and Cannon, 1997; Ganesan and Hess, 1997; Leisen and Hyman, 2001).

Marketing research has focused on trust in a particular interaction partner and on its consequences in buyer-seller relationships. Contributions mainly stem from two strands in the literature: research on relationship marketing (Ganesan, 1994; Morgan and Hunt, 1994; Doney and Cannon, 1997; Ganesan and Hess, 1997; Garbarino and Johnson, 1999) and research on electronic commerce (Jarvenpaa et al., 2000; McKnight et al., 2002a,b; Harris and Goode, 2004; Bart et al., 2005; Schlosser et al., 2006). While trust is deemed to develop over time by repeated interactions (Rempel et al., 1985; Mayer et al., 1995; Sillence et al., 2006), it also plays an important role in the first encounter between customer and company (McKnight et al., 1998; McKnight et al., 2002b; Stewart, 2003; Trifts and Häubl, 2003). Here, trust is based on customers’ first impression of a company (e.g. website design) or, if available, on more distal indicators such as size or reputation (Jarvenpaa et al., 2000; Briggs et al., 2002; McKnight et al., 2002b; Schlosser et al., 2006; Sillence et al., 2006).

Most authors in marketing research draw upon a belief-based conceptualization of trust, that is perceived trustworthiness (Gefen et al., 2003). For instance, Doney and Cannon (1997) define trust ‘as the perceived credibility and benevolence of a target of trust’ (p. 36). Trust as beliefs also prevail in the measurement of trust in electronic retailing. Though the names of the dimensions vary between the authors, their definitions and operationalizations suggest that they can be mapped onto the four dimensions ability, benevolence, integrity and predictability (McKnight et al., 2002a).

Unfortunately, a variety of measurement approaches exists because most studies use their own ad hoc scales. Furthermore, no consensus exists on whether perceived trustworthiness of online retailers is a uni- or multidimensional construct. The issue is often not explicitly explored; instead, most scales are a priori treated as unidimensional (e.g. Jarvenpaa et al., 2000; Harris and Goode, 2004; Bart et al., 2005). Some scales also tap into both beliefs and general aspects of trust (Gefen et al., 2003; for examples see Jarvenpaa et al., 2000; Bart et al., 2005).

The multidimensional approaches are hardly convincing either. McKnight et al. (2002a) present a trusting beliefs scale with three dimensions (i.e. competence, benevolence and integrity) but concede problems with discriminant validity. Moreover, they changed to a unidimensional second-order factor structure when using the items to test their trust building model (McKnight et al., 2002b). Garbarino and Lee (2003) built upon the

---

1They also integrate trust as a trait in their model. This, however, is beyond the focus of this paper and thus will not be explored further.
two dimensions benevolence and competence but had to eliminate most of the competence items because of double loadings; the remaining two items leave the competence scale very narrow in focus. Schlosser et al. (2006) applied three subscales (i.e. ability, benevolence and integrity) and found differential influences on purchase intentions. Nevertheless, they concede that the beliefs are correlated (p. 137), but neither report details on the factorial structure nor the correlations between the subscales. In addition, they introduce two completely different subscales for ability and integrity in their last experiment (Schlosser et al., 2006, study 4). In sum, the dimensionality and hence the proper measurement of perceived trustworthiness remain blurry.

**Trust effects**

Purchasing involves taking risks for consumers, that is accepting the possibility of adverse consequences (Bauer, 1960; Mitchell, 1999; Pires et al., 2004). Risk taking as a consequence of trust is inherent in most conceptualizations of trust (Mayer et al., 1995; Gefen et al., 2003) and is empirically well supported (Schlenker et al., 1973; Mayer and Davis, 1999; King-Casas et al., 2005). Risk is deemed to be even higher in online transactions because the lack of physical contact curtails consumers’ opportunities to exert control (Harris and Goode, 2004; Pires et al., 2004).

The crucial role of trust in electronic commerce is underlined by a growing body of research. Trust has been found to foster the intention to buy from a provider (Jarvenpaa et al., 2000; McKnight et al., 2002b; Konradt et al., 2003; Stewart, 2003; Trifts and Häubl, 2003; Bart et al., 2005; Schlosser et al., 2006), the intention to share personal information (McKnight et al., 2002b; Bart et al., 2005), satisfaction with a provider (Harris and Goode, 2004), the intention to follow advice (McKnight et al., 2002b) and self-reported taking of advice (Briggs et al., 2002). Trust entails also positive effects that reach beyond a single transaction: Trust enhances customers’ loyalty to both online and offline sellers (Ganesan, 1994; Morgan and Hunt, 1994; Doney and Cannon, 1997; Ganesan and Hess, 1997; Harris and Goode, 2004; Cho, 2006).

Unfortunately, most hitherto studies have relied exclusively on self-reported risk-taking – primarily the intention to buy (e.g. Bart et al., 2005). Intentions, however, cannot accurately predict whether a person really purchases something – even when assessed immediately before visiting a store (e.g. Cobb and Hoyer, 1986). To our knowledge, the study by Konradt et al. (2003) is the only one that has examined a behavioural consequence of trust. They found a moderate influence of trust on intention to buy, but a low impact of trust (compared to other variables like usability) on decision to buy from an online shop. However, it remains open whether this result is attributable to the quality of their scale – they employed the short, ad hoc scale from Jarvenpaa et al. (2000) – or to general mechanisms of trust. This highlights the importance of validating trust scales with both attitudinal and behavioural measures.

**Perceived risk**

Consumer research has demonstrated that it is not objective risk but perceived risk that matters (Bauer, 1960; Dowling and Staelin, 1994; Garbarino and Straehilevitz, 2004). Perceived risk refers to consumers’ subjective assessment of possible negative consequences that a behaviour, for example a purchase, might produce (Bauer, 1960; Dowling and Staelin, 1994; Mitchell, 1999; Jarvenpaa et al., 2000). With regard to trust, two different types of perceived risk need to be distinguished: the risk associated with a particular interaction partner (e.g. the service provider X; Jarvenpaa et al., 2000; Pires et al., 2004) and the risks perceived with that kind of

---

2Trifts and Häubl (2003) did connect their participants’ choices with monetary consequences. However, these consequences were only loose ones, because the outcome was distributed via a lottery (i.e. only one person received it), and the value of the outcome was also determined by a random procedure.
transaction in general (e.g. buying drugs; Mayer et al., 1995). Perceived risk associated with a particular interaction partner should be inversely related to the trust in the trustee: the more one trusts someone, the less risky one perceives an interaction with this partner (Jarvenpaa et al., 2000). For general perceived risk, the relationship with trust is different: the more risky a type of interaction is perceived to be, the more trust is necessary to engage in such an interaction with a particular partner (Mayer et al., 1995; Delgado-Ballester and Munuera-Alemán, 2001). If the individual does not perceive any risk at all, for example due to control mechanisms like contracts or the minor amount of money at stake, trust is not necessary to interact with this partner. In other words, general perceived risk is assumed to moderate the influence of trust on risk taking. This is also implicit in the assumption that trust is more important in electronic commerce because of its higher risks compared to traditional offline retailing (McKnight et al., 2002a; Schlosser et al., 2006).

However, empirical evidence on perceived risk as a moderator of trust and risk taking is scarce. Güürhan-Canli and Batra (2004) have addressed this issue and found the influence of the trustworthiness of a company on product evaluations to be indeed moderated by the risk of malfunction perceived for the particular type of product. Schlosser et al. (2006) demonstrate that the influence of perceived ability on intention to buy is higher when a purchase involves high (vs. low) social risk. Their conceptualization of risk, however, refers to the consequences of the purchase (i.e. intended use), not to the interaction between buyer and seller. Indirect support for the moderating influence of perceived risk is provided by Delgado-Ballester and Munuera-Alemán (2001): They found the influence of trust on customer commitment to be moderated by customer involvement. The authors argue that customer involvement is closely tied to perceived risk; this relationship is empirically well established (e.g. Laurent and Kapferer, 1985).

Research context and hypotheses

For assessing perceived trustworthiness of an online vendor, we constructed and validated a scale according to psychometric procedures and criteria (Lienert and Raatz, 1994; Nunnally and Bernstein, 1994). For the scale, we drew on the four dimensions of trustworthiness ability, benevolence, integrity and predictability (Mayer et al., 1995; McKnight et al., 1998). Online pharmacies were chosen as a suitable example of health-related electronic commerce. At the time the study was conducted, distributing pharmaceuticals via the Internet was still prohibited in Germany as well as in many other European countries. Thus, participants were expected to have no prior experience with online pharmacies and therefore to have few preconceptions. In addition, as pharmaceuticals are relevant for almost everyone and German legislation was expected to legalize online pharmacies in the near future (and has indeed done so in 2004), we assumed that participants were involved in this issue.

We derived and tested the following hypotheses on the multifarious relationships between trustworthiness, risk taking and perceived risk. The hypothesized relationships between the constructs are summarized in Figure 1.

H1. The more trustworthy an online vendor is perceived by a customer, the stronger the customer’s intention to buy from this vendor.

The intention to buy depends on the level of trust towards the provider which in turn depends on the provider’s perceived trustworthiness (Mayer et al., 1995; McKnight et al., 2002b; Bart et al., 2005; Schlosser et al., 2006).

H2. The more trustworthy an online vendor is perceived by a customer, the more likely this customer is to take an actual financial risk towards this vendor.

As a mere intention (here: to buy) does not necessarily denote that the intended behaviour (i.e. purchasing) will be exhibited by the
individual (Cobb and Hoyer, 1986), we wanted to back up our results by adding a real-life consequence of trust in the form of behavioural risk taking (Mayer et al., 1995). We opted for financial risk taking because taking a financial risk is inherent in any commercial transaction.

**H3.** The stronger a customer’s intention to buy from a provider, the more likely this customer is to take a financial risk towards this vendor.

If intention to buy and financial risk taking are both expressions of the same construct, that is risk taking, there should be a positive correlation between them (Campbell and Fiske, 1959).

**H4.** The higher the perceived risks of an online purchase in general, the higher the impact of perceived trustworthiness on intention to buy from a provider.

The more risky a transaction is perceived, the more trustworthy a provider needs to be perceived to engage in a transaction (Mayer et al., 1995; Delgado-Ballester and Munuera-Alemán, 2001). If a consumer perceives no risks at all, the intention to buy should be independent from the perceived trustworthiness of a vendor. Therefore, perceived risk is expected to moderate the relationship between trustworthiness and intention to buy.

---

**Method**

**Sample**

A total of 812 participants were recruited from the online panel (Göritz, 2007) at www.wisopanel.uni-erlangen.de, of which 52.5 per cent participated in the study. Another 275 participants were recruited via newsgroup postings and word of mouth. Overall, 687 participants completed the study. To ensure a high data quality, we excluded four groups of participants: those who spent extremely little or very much time in the pharmacy, those who produced a high rate of missing values (>10%), those who produced extreme response sets and those who were suspected of multiple submission (Birnbaum, 2004). According to these criteria, 57 cases were excluded, leaving 631 in the final sample, of which 64 per cent had been recruited from the panel. As the panellists and the external participants did not differ substantially in the control variables, we collapsed the two groups. Average age was 30.7 years (SD = 9.2), 73 per cent had attended high school, 48 per cent were employed (another 42% were students) and 93 per cent lived in Germany. Of the panel group, 41 per cent were women. Most participants were experienced.

---

3Members had been recruited for the panel in various ways, offline as well as via different websites.

4During this study, the participants were not asked to indicate their sex. Their sex had been assessed prior to this study on the occasion of signing up with the panel. Therefore, the sex of the external participants is unknown.

---

Copyright © 2008 John Wiley & Sons, Ltd.

Journal of Consumer Behaviour, Jan.–Feb. 2008

DOI: 10.1002/cb
Internet users, MD = 4 years, and indicated that they use the Internet on a daily basis (77%). Experience with online shopping was also high: 89 per cent had purchased at least one item. As the legal situation in Germany had led us to expect, 97 per cent had no experience with buying medication online. As an incentive for participation participants were promised a 5 € voucher which could be used for buying books or CDs online.

Procedure and stimuli
The study was conducted on the WWW. When entering the website, participants were informed that the goal of the study was to examine how users evaluate online pharmacies. Since at the time of the study, online pharmacies had not yet been legalized in Germany, some general information about the topic was given. Then participants answered questions on their demographics, Internet usage and perceived risk of buying medication online. On the next page, participants were informed about the limits of the simulated online pharmacy (e.g. participants’ drug orders were not actually shipped). Also, the cover story for the behavioural measure of risk taking was given: Participants were told that as a next step they would be able to visit one pharmacy randomly chosen out of two. One pharmacy would be trustworthy, the other not. In fact, however, all participants visited the same online pharmacy.

Next, participants entered the online pharmacy simulation. Participants could move around freely and as long as they liked. Users could put products into a shopping cart and initiate the ordering process. The design of the pharmacy (Figure 2) as well as products and prices were realistic, except that ordered drugs...
were not actually delivered. Moreover, the assortment was restricted to over-the-counter drugs, which do not require prescription. The participants could leave the pharmacy by a link which was displayed at the bottom of the screen. After leaving the pharmacy, they were asked about the trustworthiness of the pharmacy and their intention to buy from this pharmacy. Finally, they were given a choice to take an actual financial risk (cf. below). In order to rule out any sequence effects, the order of items pertaining to each construct was presented at random.

Measures

Before participants entered the pharmacy, perceived risk was assessed in relation to seven items referring to possible negative outcomes (cf. Appendix). Participants rated how likely they assumed these risks to occur when buying at an online pharmacy. For each item, we used a seven-point Likert-type scale, anchored by not at all and yes, definitely. Since the internal consistency was adequate, Cronbach’s α = 0.76 > 0.70, (Nunnally and Bernstein, 1994, p. 265), the average of the seven items was used as a score for perceived risks.

The following three measures were assessed after participants had left the pharmacy. For assessing perceived trustworthiness, we both coined new items and adjusted items from previous studies (Larzalere and Huston, 1980; Rempel et al., 1985; Ganesan, 1994; Doney and Cannon, 1997; Jarvenpaa et al., 2000; Kammerer, 2000; Alberternst, 2002). The items were formulated to be neutral with regard to the kind of shop. Thus, their application is not restricted to online pharmacies, but they can be used for online shops in general. For each of the four assumed dimensions of perceived trustworthiness, that is ability, benevolence, integrity and predictability, those nine items that were judged to be most relevant by three raters in a pre-test were included in the final scale. This resulted in an item pool of 4 × 9 = 36 statements. Participants were asked to rate on a seven-point scale how much they agreed (agree not at all; completely agree). These items formed the basis for the development of the trustworthiness scale, as reported in the Section ‘Results’.

For intention to buy, participants indicated the likelihood of a purchase on a seven-point Likert-type scale (Assuming this provider really existed, would you buy anything from this provider? anchored by not at all and yes, definitely).

As a behavioural measure of financial risk taking, we created a decision situation which involved a financial risk for the participants. According to the cover story, participants thought they interacted with one pharmacy randomly chosen out of two, and one of these two pharmacies would be trustworthy, the other not. At the end of the final questionnaire, participants were asked to indicate on a dichotomous scale whether they would stake their 5 € guaranteed incentive because they trusted the visited pharmacy or whether they would not stake their incentive. They were told that if they indicated that they trusted the pharmacy they had visited and this was indeed the pharmacy that was trustworthy, their 5 € incentive would be augmented to 7.50 €. If they chose to trust the pharmacy and it was the untrustworthy one, however, they would lose their 5 € incentive. By contrast, if they chose not to stake their incentive because they were not convinced of the pharmacy’s trustworthiness, they would keep the promised 5 €, regardless whether they had visited the trustworthy or untrustworthy pharmacy. Contrary to what the participants were told, all participants interacted with the same pharmacy (for fairness, all participants received at least a 5 € voucher after completing the study).

Results

Trustworthiness scale

One goal of this study was to develop a short scale for perceived trustworthiness, which can easily be administered in further studies. Therefore, we reduced the original item pool and chose the items with the best psycho-
metric quality (Lienert and Raatz, 1994; Nunnally and Bernstein, 1994). This item selection procedure was based on a randomly chosen sample of two-thirds of the original sample ($N_{SELECT} = 421$).

A principal components analysis (Varimax rotation) on the 36 items yielded three factors with an eigenvalue greater than one, explaining 52.2, 3.7 and 2.9 per cent of the variance, respectively. However, the contributions of the second and the third factor were small in terms of explained variance, and an inspection of the scree plot suggested the presence of only one factor. Therefore, we conducted a parallel analysis with 1000 random datasets (O’Connor, 2000). The results support the one-factorial solution as the eigenvalues of the second and third factor do not correspond to the recommended 95th percentile of the random data eigenvalues (O’Connor, 2000, p. 397). All items load higher than $0.46$ in the expected direction on this single factor.

In a further pruning of this item-pool, we chose the 12 best items (Table 1) in terms of item-to-total correlations ($0.86 > r > 0.75$), which were also those with the highest factor loadings ($0.87 > |r| > 0.77$). The quality of the resulting 12-item scale was cross-validated using the remaining third of the original sample as hold-out sample ($N_{VALIDATE} = 210$). Item-total correlations ($0.83 > r > 0.67$) are negligibly lower than those from the selection sample and a Cronbach’s $\alpha = 0.95$ indicated a high reliability of the scale. The scale consists of three items from each of the four postulated dimensions of perceived trustworthiness (ability, benevolence, integrity and predictability).

The factor structure of the 12-item solution was also tested in the whole sample with the help of confirmatory factor analysis using AMOS 6.0 (maximum likelihood estimation). Because the $\chi^2$ goodness-of-fit statistic has been found to underperform when using a high number of factors per indicator in

<table>
<thead>
<tr>
<th>Items</th>
<th>Sample for scale development ($N = 404$)</th>
<th>Sample for scale cross-validation ($N = 209$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor loading</td>
<td>Item-to-total correlation</td>
</tr>
<tr>
<td>Ability</td>
<td>This provider is very competent</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>This provider is able to fully satisfy its customers</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>One can expect good advice from this provider</td>
<td>0.86</td>
</tr>
<tr>
<td>Benevolence</td>
<td>This provider is genuinely interested in its customers’ welfare</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>This provider puts customers’ interests first</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>If problems arise, one can expect to be treated fairly by this provider</td>
<td>0.77</td>
</tr>
<tr>
<td>Integrity</td>
<td>I am happy with the standards by which this provider is operating</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>This provider operates scrupulously</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>You can believe the statements of this provider</td>
<td>0.87</td>
</tr>
<tr>
<td>Predictability</td>
<td>This provider’s methods of operation are unclear$^\dagger$</td>
<td>−0.77</td>
</tr>
<tr>
<td></td>
<td>This provider keeps its promises</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>I would rely on advice from this provider</td>
<td>0.80</td>
</tr>
</tbody>
</table>

Table 1. Items for measuring perceived trustworthiness

Note: For the German wording of the items please contact the first author.
Based on one-factorial solution.
Reverse coded.
combination with small sample sizes (Marsh et al., 1998), which is the situation in our study, we followed the recommendations by Hu and Bentler (1999) on robust indexes and cut-off criteria. The respective fit indexes demonstrate that the model fits the data well (recommended cut-off criteria in parentheses): CFI = 0.993 (>0.95), RMSEA = 0.035 (<0.05) and SRMR = 0.025 (<0.09); item loadings are between |0.74| and |0.85| (all ps < 0.001) and in the expected direction.

The unidimensionality of the solution was backed-up by the results from testing a competing four-dimensional model. For this, the four originally postulated dimensions were introduced as latent variables and each of the 12 items from the reduced item pool loaded on the dimension for which it had originally been developed. This resulted in a model with three items for each of the four factors ability, benevolence, integrity and predictability. The fit indexes are only marginally different from the one-dimensional model: CFI = 0.994 (>0.95), RMSEA = 0.036 (<0.05) and SRMR = 0.024 (<0.09); item loadings are between |0.75| and |0.85| (all ps < 0.001) and in the expected direction. However, the correlations between the four factors are all >0.97, thus indicating a lack of discriminant validity of the four-dimensional model.

These findings strongly support the one-dimensional model, which is also more parsimonious. Hence, the mean of these 12 items was calculated for each individual in the whole sample and used in the remaining analyses as a score for trustworthiness.

**Hypotheses tests**

**Intention to buy**

Perceived trustworthiness and intention to buy correlated positively, \( r = 0.74, p < 0.001, N = 621.5 \)
Therefore, Hypothesis 1 is supported: the higher the perceived trustworthi-

---

5 All hypotheses were tested one-tailed at an \( \alpha \)-level of 0.05 on the whole sample.
ness are fulfilled: (1) regressing perceived trustworthiness on perceived risk yielded a negative effect, $\beta = -0.31$, $p < 0.001$, (2) regressing intention to buy on perceived risk also yielded a negative effect, $\beta = -0.34$, $p < 0.001$, (3) when regressing intention to buy on both perceived risk and perceived trustworthiness, the coefficient for perceived trustworthiness is significant (Table 2, Step 1). Finally, (4) the regression coefficient for perceived risk is lower in regression (3) than in regression (2) ($\beta = -0.13$ vs. $\beta = -0.34$). In sum, the effect that an increase in perceived risk lowers the intention to buy is partially mediated by reducing the perceived trustworthiness of the vendor. Moreover, the impact of perceived risk on intention to buy is smaller than the impact of perceived trustworthiness (Table 2).

Control variables

To control for selection bias, we compared the panelists to the external participants with regard to perceived risk, behavioural risk taking, perceived trustworthiness and intention to buy; no significant difference was found. Furthermore, neither time spent in the pharmacy nor the number of pages visited correlated with the variables used for testing the hypotheses.

Table 2. Hierarchical regression analysis for predicting intention to buy ($N = 616$)

<table>
<thead>
<tr>
<th>Variable</th>
<th>$B$</th>
<th>SE $B$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived trustworthiness (PTW)</td>
<td>1.10</td>
<td>0.05</td>
<td>0.70***</td>
</tr>
<tr>
<td>Perceived risk (PR)</td>
<td>-0.23</td>
<td>0.05</td>
<td>-0.13***</td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTW</td>
<td>1.10</td>
<td>0.05</td>
<td>0.70***</td>
</tr>
<tr>
<td>PR</td>
<td>-0.24</td>
<td>0.05</td>
<td>-0.13***</td>
</tr>
<tr>
<td>PTW $\times$ PR</td>
<td>0.02</td>
<td>0.04</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Note: $R^2 = 0.56$, $p < 0.001$, for Step 1; $\Delta R^2 < 0.01$, n.s., for Step 2.***$p < 0.001.$

Discussion

One achievement of this study is the development and validation of a scale to measure the perceived trustworthiness of online shops. The trustworthiness scale is sound with regard to psychometric criteria. The scale’s reliability is excellent and it has proven its prognostic validity, not only for the self-reported intention to buy but for actual financial risk taking as well. A laboratory study that used existing online pharmacies provides further evidence for the scale’s validity: perceived trustworthiness as measured by this scale predicted intention to buy as well as participants’ choice between competing providers (Büttner et al., 2006). Moreover, the scale can also be used in offline retailing because the wording of the items is not specific to an online context. Preliminary results show that the psychometric quality of the scale holds up when applied in brick and mortar retailing.

In the course of the scale’s development, the postulated dimensions of perceived trustworthiness (i.e. ability, benevolence, integrity and predictability) were not confirmed to be independent factors. Previous research has yielded mixed results with regard to the factor structure of trustworthiness. Most studies have failed to detect more than one dimension (e.g. Larzalere and Huston, 1980; Doney and Cannon, 1997). By contrast, Ganesan and Hess (1997) were able to establish a scale with two factors. The same applies for the study by Garbarino and Lee (2003); their analysis, however, reveals problems with establishing independent factors.

To explain these contradictory findings with regard to the underlying factor structure of trustworthiness, we propose to take the duration of a relationship into account. As in a first encounter, there is no or little knowledge about a partner, an evaluation of the partner’s trustworthiness can only be a global and therefore hardly differentiated judgment. As a
relationship matures through repeated interactions, the picture of the partner becomes more multifaceted and, for instance, a certain vendor might be judged as benevolent but incompetent at the same time. Indeed, McKnight et al. (2002a, p. 348), who failed to convincingly establish a four dimensional solution for their trusting beliefs subscale, examined a first encounter between participants and a new website. In our study, not only was the vendor new but also the whole domain of buying drugs online. On the contrary, in the study by Ganesan and Hess (1997) the minimum duration of relationship between buyer and seller was 4 months; and 89 per cent had interacted with each other for at least 1 year. The participants in the study by Garbarino and Lee (2003) rated well-known online retailers. Further support comes from research on social perception, which indicates that a halo effect is very likely to occur if the person to be evaluated is unknown to the rater (Kozlowski et al., 1986). Thus, further research needs to address the impact of the stage of the relationship with a vendor on the dimensionality of perceived trustworthiness. The balanced contribution of the four postulated dimensions of perceived trustworthiness (i.e. ability, benevolence, integrity and predictability) to the developed trustworthiness scale (i.e. each dimension is represented with three items) facilitates further study of this issue.

In line with recent research on electronic commerce, the study demonstrates that perceiving an online vendor as trustworthy is crucial for consumers’ decision to buy from this vendor. It extends current research by replicating this finding both for another domain and for a more heterogeneous sample than those previously investigated (Jarvenpaa et al., 2000; McKnight et al., 2002b; Stewart, 2003; Trifts and Häubl, 2003; Schlosser et al., 2006). Moreover, the study demonstrates an influence of perceived trustworthiness on purchase decision not only based on self-reports, such as intention to buy, but based on actual financial risk taking as well. Hence, trust can indeed be assumed to play a crucial role in real-life purchases.

Furthermore, the study sheds light on a blind spot in trust research: the assumed moderating role of perceived risk for the relationship between perceived trustworthiness and risk taking. Our data do not support the assumption of a moderating role of perceived risk but suggest that the influence of perceived risks on intention to buy is partially mediated by perceived trustworthiness. However, the mediating role of perceived trustworthiness was not an a priori hypothesis in this study; therefore, it should be explicitly tested in further studies. Moreover, in light of Schlosser et al. (2006), Gurhan-Canli and Batra (2004) and Delgado-Ballester and Munuera-Alemán (2001), other plausible explanations should be considered before abandoning a possible moderating role of perceived risk.

An alternative explanation for this study’s disconfirmation of the moderating influence of perceived risk might be a domain-specific ceiling effect: individual differences in perceived risk might have been drowned out by the high perceived risk in the medical domain in general. If this is the case, the moderator effect should resurface in differences in the strength of the relationship between perceived trustworthiness and purchase behaviour for different domains of e-commerce (e.g. low for books, high for drugs). Indeed, the correlation of perceived trustworthiness and intention to buy found in our study exceeds effect sizes found in less risky domains (Jarvenpaa et al., 2000; McKnight et al., 2002b; Konradt et al., 2003; Stewart, 2003; Schlosser et al., 2006), although not in all studies (Trifts and Häubl, 2003; Bart et al., 2005). As these studies are comparable only in part and the online shops examined differ by more features than only perceived risk, clarification must await further studies.

By simulating a pharmacy, it was ensured that all participants interacted with the website for the first time. As a downside of not using a real online shop with actual shop visitors, however, participants might have evaluated the simulated pharmacy differently than visitors of an actual pharmacy. Our data cleaning
and a thorough inspection of the server’s log files have ensured that conspicuous datasets were discarded.

The limitation of our study to one sector and one country bears the danger that the findings are context-specific because online consumer behaviour differs between sectors and countries (Harris and Goode, 2004; Bart et al., 2005). For instance, in the medical domain more players than merely a consumer and a provider are involved (e.g., physicians, health insurers). However, the simulated pharmacy together with the fact that a market did not yet exist allowed us to abstract from these and other idiosyncrasies encountered in more mature markets (e.g. established retailers) and to focus on the basic processes of online trust. Congruent with this reasoning, the impact of perceived trustworthiness has also been found in studies from different countries using both fictitious and existing online shops from various sectors (Jarvenpaa et al., 2000; McKnight et al., 2002b; Stewart, 2003; Bart et al., 2005; Büttner et al., 2006; Schlosser et al., 2006). This strengthens the generalizability of the outcome and indicates that the peculiarities of other sectors and countries might impact the levels of perceived risk and perceived trustworthiness, but do not alter the basic processes. To further clarify this issue, we recommend that future studies always include control measures for the level of risk perceived in the domain under scrutiny.

The findings from this study have two major implications for the management of online shops. First, investing into building and communicating trustworthiness is worthwhile as it is crucial for attracting customers. As argued above, the generalizability of this statement is supported by results from other countries and sectors. Second, the key to success lies in the hands of the particular online shop because the impact of the shop’s trustworthiness is higher than the impact of general perceived risk, which is beyond the influence of the retailer. Establishing trustworthiness as a component of the company’s image can result into a competitive advantage – especially in high-risk domains.

Overall, the results imply that online shops should put effort into the systematic management of trust. The developed perceived trustworthiness scale can be a valuable tool for this purpose, for instance when comparing one’s own trustworthiness to those of competitors or when (re-)designing an online shop.

Acknowledgements

This work was supported by the University of Erlangen-Nürnberg postdoctoral scholarship (HWP) to the second author. We thank Klaus Moser for his helpful input during the planning and fielding of this study, and the two anonymous reviewers for their helpful comments.

Biographical notes

Oliver Büttner is a PhD candidate in marketing at the Georg-August-Universität Göttingen, Germany. His research activities revolve around shopping behaviour in electronic and in brick & mortar retail environments. Currently, he is working on methods for field research at the point of purchase.

PD Dr. Anja Göritz (www.goeritz.net) researches on Web-based data collection, especially online panels, mood induction, and incentive effects. In 2000 she built and has since maintained Germany’s first university-based online panel with more than 5000 panelists (www.wisopanel.uni-erlangen.de). She programmed serveral platforms and tools for Web-based data collection (e.g. www.goeritz.net/brmic and www.goeritz.net/panelware). She has regularly been an instructor in the Advanced Training Institute “Performing Web-Based Research” of the American Psychological Association.

References


O’Connor BP. 2000. SPSS and SAS programs for determining the number of components using


Appendix: Items for perceived risk (for the wording of the items in German please contact the first author)

1. Information about your illnesses and/or your use of medication falls into the hands of unauthorized people.
2. You are wrong or inadequately advised.
3. You receive counterfeit, adulterated or expired medication.
4. Your bank account or credit card is charged improperly.
5. Medication is damaged during delivery (e.g. breakage or interruption of cold chain).
6. Your personal data (e.g. e-mail or postal address) are used for other purposes than indicated (e.g. spam).
7. The promised medication is delivered too late or not at all.