

E-COMMERCE CREDIBILITY IN SAUDI ARABIA: CONSUMER PERSPECTIVE

Shahad Khoja¹ and Khalid Al-Omar²

^{1,2}Department of Information Systems, King Abdulaziz University, Jeddah, Saudi Arabia

ABSTRACT

The purpose of this paper is to investigate the e-commerce credibility factors affecting the perception of users in Saudi Arabia and, moreover, to investigate whether the variation of credibility factors in Saudi Arabian e-commerce websites influence users' performance. Website credibility, which refers to the believability of the website and its content, plays an important role in consumers' successful online shopping experience and satisfaction. This investigation is conducted by employing two credibility evaluation methods: heuristic evaluation and performance measurement. This study adopts Fogg's 10 Stanford credibility guidelines as a starting point for the heuristic evaluation. In the performance measurement method, two measurements are used: the amount of time needed to finish the task and the total number of clicks taken to finish the task. A frequency analysis of the comments and a one-way ANOVA test are used to establish the results. Three e-commerce websites in Saudi Arabia are selected. The findings show that Fogg's 10 Stanford credibility guidelines can be implemented in the Saudi Arabian e-commerce context with minor modifications and expansions by adding reputation, endorsement, security, and service diversity guidelines. Another important finding is that professional website design plays a vital role in users' first impression of websites, while usability is the most important credibility factor investigated used to evaluate the credulity of an e-commerce website. Lastly, the results of this study indicate a relationship between the e-commerce credibility level and users' performance. This paper contributes to the literature by providing a set of credibility guidelines associated with specific criteria, which can be assessed to improve the future of e-commerce in Saudi Arabia.

KEYWORDS

Credibility, e-commerce, heuristic evaluation, performance measurement.

1. INTRODUCTION

Persuasive technologies, unlike other technologies, inherently transformative. They comprise interactive systems designed to change attitudes and behaviours based on the user's current behaviours or attitudes. Credibility expresses and supports persuasion. Many research strategies have been employed to apply credibility [1], such as those investigating the elements that affect the credibility of interface design [2–4] and those that have indicated the effects of individual differences on the user perception of credibility [5–7].

Elements that increase the credibility of websites have become a subject of interest to most researchers. Numerous studies have been conducted to reveal credibility elements, including appearance and aesthetics [4, 8], ease of navigation and use [9], citations [10], the number of advertisements [11, 12], and the relevance of the website content to its advertised product [13]. Another subject of interest is the first impression. Peracchio and Luna (2006) indicated that 80% of website visitors spend only a few seconds "inspecting" a webpage before leaving and moving

to another. Consequently, when the user is looking at the website, the first impression formed in the first few seconds, which in turn will affect the user's decision to either stay or leave the website. The first impression of the user navigating the website is an essential key to success for many organisations operating digitally, particularly e-commerce businesses, because websites are the primary interface that represents the organisation and is the means through which the customer can interact with the organisation (such as accessing specific information, buying and selling items, or receiving customer service). On the other hand, persuasion technologies are becoming ubiquitous, being used in various areas, including marketing, tourism, health and customer services. At the same time, most of these technologies are designed for an English-speaking audience. Now, with the remarkable growth of the Internet, more and more users around the globe, and increasing customer demands, the trend is slowly changing such that web developers are taking into consideration different cultural norms. Because persuasion technology in the real world associated with cultural standards, it is essential for developers of persuasion systems to consider cultural norms when designing and developing systems for the non-English-speaking market. Khaled et al. (2006) pointed out that most existing studies have focused on the individualistic Western culture and that persuasion strategies may demonstrate varying degrees of effectiveness in different, collectivist cultures.

2. CREDIBILITY

Credibility is a complex concept [6]. It can be defined best as believability [1]. A credible person is a believable person, credible information is trustworthy information, and a credible website is a believable website. Two fundamental factors are considered to be related to credibility: trustworthiness and expertise [16]. The former falls into the concept of reliability [1], while the latter is related to the knowledge and skills of the source [17].

Fogg identified four types or channels of credibility: presumed credibility, reputed credibility, surface credibility, and experienced credibility [17].

- Presumed credibility describes the degree to which a perceiver believes in someone/something based on the perceiver's general assumptions. This type is highly affected by the perceiver's cultural background.
- Reputed credibility describes the degree to which a perceiver believes in someone/something based on a third party.
- Surface credibility describes the degree to which a perceiver believes in someone/something based on simple inspection.
- Experienced credibility describes the degree to which a perceiver believes in someone/something based on first-hand experience.

However, these types are not mutually exclusive: a person's perception of credibility in one context can evolve into perceived credibility in another context [2]. Moreover, credibility can be defined in terms of the characteristics of the information [18], the features of the content [4], and the delivery media [19]. Consequently, some studies have provided insight into the credibility of information and content, such as scholarly information [6]. Other research has focused on credibility factors related to computer-based media [2, 4, 11]. More importantly, credibility has a

significant effect on users' behaviour. Huang (2010) indicates that a highly credible and usable website implies better users' performance and vice versa.

3. PROBLEM STATEMENT

In Saudi Arabia, like most fast-developing countries, small and medium enterprises, as well as large corporations, have adopted e-commerce models to boost their business growth. However, even with the increasing adoption of e-commerce, one study has indicated lack of trust as a growing challenge in the sector [40]. Despite the substantial global growth in e-commerce, the adoption of e-commerce in Saudi Arabia has been limited [42]. Research has indicated a low adoption rate of e-commerce from businesses and customers alike [41], and even though the number of Internet users in Saudi Arabia has been growing substantially, the majority of businesses in Saudi Arabia have not adopted e-commerce into their business strategies [42]. The reason for weak e-commerce adoption includes many barriers, such as the lack of privacy and security [42, 39, 41], issues related to consumers' increasing awareness [21], and the description and value of products [39]. These related issues have all been identified under the concept of credibility in previous studies, and it is thus necessary for stakeholders to come up with policies or guidelines that enhance the trustworthiness between business and consumers, as well as understand consumer perception and behaviour in terms of credibility. Such an adoption would increase the number of people who rely on e-commerce for their business dealings. With time, the increased trust in e-commerce businesses would foster growth of the transition from offline to online business models in Saudi Arabia. The revolution of the entire Saudi Arabia economy would enhance the overall efficacy and lower the cost of doing business. Consequently, an increase in the credibility of e-commerce websites would positively influence some economic growth objectives of Saudi Arabia's 2030 Vision [21], such as an influx in investment and the establishment of new channels for local industry. The study of factors affecting the credibility of e-commerce in Saudi Arabia thus seeks to understand consumer perception and business-related elements that may influence credibility levels.

4. RESEARCH QUESTION

- A. Are the credibility factors mentioned in previous studies applicable to Saudi Arabia?
- B. Do the credibility differences in e-commerce websites influence users' performance?

5. WEBSITE SELECTION

Among the variety of e-commerce websites that exist, local e-commerce in Saudi Arabia is chosen for a number of reasons. First, local e-commerce is the closest level to the local user. Second, previous research has found evidence of challenges faced by local e-commerce websites, as Saudi Arabia is an emerging country in the field of e-commerce [21]. Taking these conditions into consideration, three local e-commerce websites were selected: e-commerce website 1, Namshi.com; e-commerce website 2, warqat.com; and e-commerce website 3, akoonstore.com. Local e-commerce websites that provide a wide range of credibility were selected: The aim was not to select the most credible websites but rather to present websites with variations in credibility to most effectively determine the relationship between website credibility level and users' performance as well as to identify a variety of credibility factors of e-commerce websites.

6. CREDIBILITY EVALUATION METHOD

6.1. Study 1: Heuristic Evaluation Method

Of the two evaluation methods used, the heuristic evaluation method is the easier, quicker, and more effective in terms of identifying potential problems, and it has been used in numerous studies [6, 7, 20, 22]. The heuristic evaluation method requires the evaluators to identify the interface problems against a set of heuristics, guidelines or design principles. This method's popularity is reflected in its cost-effectiveness. Additionally, the heuristic evaluation can be employed by both experts and novices, with evaluations by experts appropriately dubbed "expert evaluations." Furthermore, while the method can be conducted by a single evaluator, its effectiveness and improvement increase with the number of evaluators [23]. Fogg's 10 Stanford credibility guidelines [24] has been widely used in evaluating the credibility of websites, and a number of studies have tested it and confirm its usefulness [7, 20, 22]. These guidelines each describe a specific element of a website, and when using them in credibility evaluations, they help the evaluator focus on that specific element of website design to identify the credibility problem. Therefore, this study applies these guidelines as a starting point to evaluate the credibility of local e-commerce websites. However, Fogg's 10 Stanford credibility guidelines are too general for a heuristic evaluation, and therefore, criteria associated with each credibility guideline must be established. Table 1 shows the credibility guidelines and their associated criteria from relevant studies. The heuristic evaluation was conducted using 24 criteria on a five-point Likert scale. The instruments used are the selected websites – e-commerce1, e-commerce2, and e-commerce3 – and the heuristic evaluation sheet.

Table 1. Credibility guidelines with associated criteria

Credibility guideline	Associated criteria	Relevant studies
G1. Easy verification of information accuracy	(1) The site provides links to third-party sites for more information (e.g., when explaining policies, terms, and conditions). (2) The domain name is compatible with the company name. (3) The information is complete and accurate.	[6, 9, 10, 22, 25, 26]
G2. Demonstration of the existence of the real organisation behind the website	(4) The site displays the company's membership with the Chamber of Commerce. (5) The site lists the physical address of the company.	[6, 10, 22, 27, 28]
G3. Demonstration of the expertise in the organisation	(6) The site identifies and displays any awards it has. (7) The site identifies and displays any affiliations with other companies (e.g., delivery companies). (8) The company displays its credentials (e.g., sold out items).	[6, 9, 10, 22, 25, 26, 27, 28]

G4. Demonstration of the honesty and trustworthiness of the people behind the website	(9) The site contains an "About us" page that features information about the company's value and history. (10) The site displays the names of the people responsible for or in charge of the company.	[9, 10, 11, 25, 26, 22, 27, 28]
G5. Easily accessible contact information or features	(11) The site provides a help function, such as an "Ask a Question" or "Live Chat" feature. (12) The site provides a "Contact Us" page or an e-mail address for enquiries. (13) The site provides a toll-free phone number.	[10, 22, 26, 27]
G6. Professional design of the website	(14) The site is designed to look professional. (15) The logo and any images are high quality. (16) The site pages, design and content are presented consistently.	[2, 4, 6, 8, 9, 10, 22, 25, 27, 28]
G7. Ease of use and usefulness of the website	(17) The site is available in the user's language. (18) The site loads in an appropriate amount of time. (19) The site provides error messages that use plain language.	[9, 10, 25, 26, 22, 27, 28]
G8. Frequency of website updates	(20) The site is updated frequently.	[10, 22, 25, 26, 27]
G9. Limited advertisement material	(21) The site is free from advertisements.	[9, 10, 11, 22, 25, 27, 28]
G10. Avoidance of errors	(22) The site provides error messages when errors are detected. (23) The site is free from typographical errors. (24) Each link in the website connects to the relevant page.	[10, 22, 31, 27]

6.2. Study 2: Performance Measurement Method

Another evaluation method used in this study is the performance measurement method. Its usefulness has been presented by Huang (2010), whose heuristic evaluation and performance measurement suggested a relationship between participants' overall perception of credibility and usability and their performance. In general, in a performance measurement test, the participant is required to perform a set of practical tasks using the website or system. The participant's task performance is measured based on a number of "performance measurement criteria." This method can indicate a user's interaction levels when performing a set of tasks using the system. For example, Park and Hwan Lim (1999) implemented a performance measurement test to assess the capability of users when using a system. A number of performance criteria are measured, such as the number of errors encountered, the time spent using help or documentation, and the frequency of complaints. Matera et al. (2002) combined two evaluation methods – usability inspection and performance measurement – and concluded that these two methods are complementary and can

be used together. Ahmed et al. (2006) used both the heuristic evaluation and performance measurement methods, additionally concluding that combining these methods reveals not only the problems with the interface but also users' performance when confronted with the interface problem.

The performance measurements used are the total time needed to complete each task and the number of steps taken to finish the task. These performance measurements are included in the performance sheet and used by the examiner to assess the participants' performance. The task sheet includes a set of tasks developed by the examiner to be performed by the participants. Though task sheets can contain many tasks, in the current study, only a small number of tasks have been chosen out of respect for the participants' time.

Furthermore, the tasks in the current study are selected based on the users' perspective, not that of the experts. Thus, the tasks are based on activities users frequently engage in. The tasks are divided into two parts: the first task, labelled "first impression," which simply requires the participants to inspect the website generally, and specific tasks, which require the participants to follow specific task directions. Each task sheet contains the same tasks with a little variation to address the differences in the three e-commerce websites. The following are some example of the first impression task and specific tasks:

First impression task 0: "Navigate the website."

Specific task 1: "Search for (specific product), select a size and colour, add it to the cart, and proceed to the checkout."

Furthermore, to address the first research question, the participants are requested to answer two open-ended questions after each task practicing: the credibility factors positively affecting user perception, and credibility factors negatively affecting user perception. The main purpose of repeating the same two questions after each task is to ensure that participants' perception is captured throughout the website navigation journey and not just based on the first impression. Furthermore, this approach helps uncover the first-impression credibility factors, which are responsible for either completing the process of navigating the website or leaving it [14]. The e-commerce websites studied are the same ones used in the heuristic evaluation study. The main objective is to find the relationship between the participants' performance and the website credibility level. Therefore, the instruments used to obtain the results are the task sheet, the selected e-commerce websites, and the performance measurements.

7. PARTICIPANTS

In terms of choosing the number of participants in the study, Nielsen and Molich (1990) found that three participants can identify half of the main problems. Virzi (1992) found that between four to five participants can detect up to 80% of the problems, while 10 participants can detect up to 90% of the problems. Dumas and Redish (1990) argued that six to twelve participants is a good number to detect information, and additional participants are less likely to identify any new information. Moreover, research time, budget, and the importance of statistical significance must be taken into account. Based on these considerations, 36 participants were selected for the current study and asked to evaluate each target e-commerce website, with 12 participants assigned to each e-commerce website. Also, to measure the participants' performance, the same 12 participants who participated in the heuristic evaluation study were asked to participate in the

performance measurement study. Some participants were recruited from personal networks, and others were recruited from public places. Demographic information, such as age, gender, education level, number of purchases in the last three months and the area of study, were also collected.

8. FINDINGS AND DISCUSSION

8.1. Study 1: Heuristic Evaluation

The one-way ANOVA test used to analyse the data, it can comparatively analyse the overall credibility evaluation among the three e-commerce websites, as shown in Table 2. The results show that a significant difference ($F=22.91$, $P=0.00$) exists in the overall participant perception regarding the credibility of the three e-commerce websites. A lower mean indicates a worse overall assessment, while a higher mean indicates a better overall assessment. Accordingly, based on the information displayed in Table 3, e-commerce3 has the worst overall evaluation, with a mean of 2.78 and a standard deviation of 0.876. E-commerce2 is placed next, with a mean of 3.30 and a standard deviation of 0.944. E-commerce1 has the best overall evaluation, with a mean of 3.60 and a standard deviation of 0.595.

Table 2. One-way ANOVA test (e-commerce comparison)

Overall credibility					
e-commerce1		e-commerce2		e-commerce3	
Mean	SD	Mean	SD	Mean	SD
3.60	0.595	3.30	0.944	2.78	0.876
Significance			F= 22.91, P=0.00		

Table 3. The mean and std. deviation for each credibility criterion in the three e-commerce websites

	e-commerce1		e-commerce2		e-commerce3	
Criteria	Mean	SD	Mean	SD	Mean	SD
1	3.17	0.577	3.00	0.853	2.00	0.953
2	4.33	0.651	4.17	0.937	3.75	0.866
3	4.08	0.669	3.42	0.793	3.08	0.996
4	3.00	0.000	2.75	0.965	2.83	0.389
5	2.75	0.622	2.42	0.669	3.08	1.379
6	2.92	0.515	2.58	0.900	2.00	0.853
7	3.08	0.669	3.67	1.155	3.58	1.084
8	3.50	0.674	2.83	0.937	2.00	0.739
9	3.50	0.905	3.17	1.030	3.58	0.793
10	2.50	0.798	2.33	0.888	2.08	0.900
11	2.83	0.718	3.25	1.288	1.83	1.030
12	4.00	0.603	3.92	0.669	3.75	1.288
13	3.17	0.577	3.17	0.937	2.42	0.900

14	4.25	0.452	3.58	1.311	1.83	1.115
15	4.58	0.515	4.00	0.953	1.58	0.669
16	4.33	0.492	3.83	0.937	1.67	0.985
17	4.50	0.522	3.92	0.996	2.33	1.303
18	4.17	0.577	3.58	0.996	3.75	0.866
19	3.33	0.492	2.92	0.900	3.42	0.669
20	3.25	0.452	3.00	0.426	2.92	0.289
21	3.83	0.835	3.83	1.030	3.92	0.289
22	3.33	0.492	3.42	0.793	3.33	0.985
23	3.83	0.718	3.00	1.206	2.92	0.793
24	4.25	0.754	3.42	1.084	3.08	0.900
Total	3.60	0.595	3.30	0.944	2.78	0.876

8.2. Study 2: Performance Measurement

A one-way ANOVA test was also used to analyse the results of the performance measurement test. Tables 4, 5 and 6 show the results of users' task performance for e-commerce websites 1, 2 and 3, respectively.

Table 4. Mean and standard deviation of users' performance in e-commerce1

e-commerce 1	Number of clicks to complete the task		Time spent on the task	
	Mean	SD	Mean	SD
Task 1	11.17	4.30	1.07	0.332
Task 2	10.30	9.95	1.03	0.32
Task 3	13.75	8.83	1.72	0.46
Total	35.25	23.12	3.81	1.11

Table 5. Mean and standard deviation of users' performance in e-commerce2

e-commerce 2	Number of clicks to complete the task		Time spent on the task	
	Mean	SD	Mean	SD
Task 1	25.42	23.82	1.84	0.60
Task 2	10.80	7.78	1.57	0.47
Task 3	32.33	15.41	4.42	1.02
Total	68.50	74.01	6.82	2.09

Table 6. Mean and standard deviation of users' performance in e-commerce3

e-commerce 3	Number of clicks to complete the task		Time spent on the task	
	Mean	SD	Mean	SD
Task 1	25.42	23.82	2.42	1.00
Task 2	11	8.25	1.63	0.661
Task 3	31.67	15.02	3.44	1.15
Total	68.08	47.09	7.50	2.82

In addition, Table 7 shows the participants' overall performance results for the three e-commerce websites. The results indicate a significant difference ($F=4.490$, $P=0.019$) in participants' performance in terms of the number of clicks required to complete tasks. In detail, the participants using e-commerce1 required the lowest number of clicks for all tasks, with the mean of the overall number of clicks required to complete the tasks 35.25 and a standard deviation of 23.12 (Table 4). In second place is e-commerce3, with a mean of 68.08 and a standard deviation of 47.09 (Table 6). Finally, e-commerce2 necessitated the most clicks, with a mean of 68.50 and a standard deviation of 74.01 (Table 5).

Table 7. One-way ANOVA test (comparison of users' performance)

Number of clicks					
e-commerce1		e-commerce2		e-commerce3	
Mean	SD	Mean	SD	Mean	SD
35.25	23.12	68.50	74.01	68.08	47.09
Significance $F=4.490$, $P=0.019$					
Total time					
e-commerce1		e-commerce2		e-commerce3	
Mean	SD	Mean	SD	Mean	SD
3.81	1.11	6.82	2.09	7.50	2.82
Significance $F=17.173$, $P=0.000$					

On the other hand, the results indicate a significant difference ($F=17.173$, $P=0.000$) in the participants' performance in terms of the time required to complete all tasks. As presented in Table 7, the participants using e-commerce3 took a longer time to complete the tasks, with the mean of 7.50 and a standard deviation of 2.82 (Table 6), followed by e-commerce2 (Table 5). Participants using e-commerce1 required the smallest amount to time to complete the tasks (Table 4).

Moreover, all common factors of credibility that emerged from the open-ended questions are presented in Table 8. First, each comment is coded according to its content. The study generated 314 comments regarding the credibility of the e-commerce websites. Some of the comments were brief, while others were lengthy. Therefore, the lengthy comments are assigned more than one code. For example, the comment below would be coded under two factors: completeness and accuracy of product information and high-quality pictures and logos.
 "The product description is very bad and pictures as well."

After coding each comment, the frequency for each comment (i.e., how often the specific issue was mentioned) was also calculated. For example, the completeness and accuracy of product information is mentioned 44 times among the 314 comments – 14.01% of the time. This frequency represents an indication of what criteria people use to make credibility judgements of e-commerce websites.

Table 8. Users' comment

Criteria	Percentage
Information Accuracy	20.38%
(1) The return and exchange policies are complete and accurate.	1.27%
(2) The payment methods and procedure are complete and accurate.	0.96%
(3) The site provides links to third-party sites for more information (e.g., when explaining policies, terms and conditions).	0.64%
(4) The domain name is compatible with the company name (e.g., www.domin-name.com)	0%
(5) The shipment methods and related details are complete and accurate.	3.50%
(6) The product information is complete and accurate.	14.01%
The organisation behind the website	3.50%
(7) The site uses different channels (e.g., Instagram) to represent its services or to market itself.	0.96%
(8) The site represents the company's membership with the chamber of commerce.	1.59%
(9) The site lists the physical address of the company (e.g., Saudi Arabia, Jeddah, Abdullmaqsoud Khoja, P.O. Box XXXX)	0.64%
(10) The site has a dedicated mobile-phone application.	0.32%
Expertise in the organisation	0.32%
(11) The site identifies and displays any awards it has.	0%
(12) The site identifies and displays its relationship to other companies, such as its delivery company (e.g., Aramex).	0.32%
(13) The company displays its credentials, such as (sold out item).	0%
Demonstration of the honesty and trustworthiness of the people behind the website	1.91%
(14) The site contains an "About us" page that features information about the company's value and history.	1.27%
(15) The site displays the names of the people responsible for or in charge of the company.	0.64%
Easily accessible contact information or features	3.18%
(16) The site provides a help function, such as an "Ask a Question" or "Live Chat" feature.	1.27%
(17) The site provides a "Contact Us" page or an e-mail address for enquiries.	0.96%
(18) The site provides a toll-free phone number.	0.64%
(19) The site provides its social media pages as a means to contact the organisation.	0.32%
Professional design of the website	14.97%
(20) The site is designed to look professional.	5.73%
(21) The logo and any images are high quality.	7.96%
(22) The site pages, design and content are presented consistently.	1.27%
Frequency of website updates	0%

(23) The site is updated frequently.	0%
Advertisement material	0%
(24) The site is free from advertisements.	0%
Ease of use and usefulness of the website	29.62%
(25) [a variety of different usability criteria]	29.62%
Reputation	6.37%
(26) The products have a good reputation.	0.96%
(27) The website has a good reputation.	5.41%
Security	1.59%
(28) The site is secure.	0.64%
(29) The site provides a privacy policy.	0.32%
(30) The site uses secure payment methods such as PayPal.	0.64%
Endorsement	6.69%
(31) The site features customer reviews.	4.78%
(32) The site features customer reviews on its social media sites.	1.27%
(33) Users recommend this site.	0.64%
Service diversity	11.46%
(34) The site offers different payment methods.	9.23%
(35) The site offers different shipment methods.	0.64%
(36) The site offers different services (shipment tracking, related products, etc.).	1.59%

The first credibility guideline, information accuracy, refers to all related comments in which people expressed confirmation or doubt about the product information on the site. Table 8 shows that the response rate regarding product information and accuracy is 14.01%, and the rate for the accuracy of information pertaining to shipment methods is 3.50%. Conversely, the compatibility between the domain name and the company name received no responses or interest from the participants. Unlike the previous study [12], where the users response is indicated. The main weakness stems from the survey methodology used in the previous study that it doesn't capture the actual users' behavior.

The second credibility guideline, the organisation behind the website, represents 3.50% of the total comments. New credibility criteria not previously mentioned were added to the second credibility guideline: 1) using different channels to represent the website's services or to market it and 2) the existence of a dedicated mobile phone application. Although the number of responses is low, these questions show that social media and mobile phone applications play some role in e-commerce websites' credibility.

The third and fourth credibility guidelines, highlighting expertise in the organisation and demonstrating trustworthiness, represent only 0.32% (only one comment out of 314 comments) and 1.91%, respectively. This finding runs contrary to those of previous studies, which have suggested that expertise and trustworthiness are critical components of credibility and that users make assessments based on these two components to arrive at an overall credibility judgement. A possible explanation for this result might be that most studies are conducted in the area of scholarly information [6] on the web. It is possible, therefore, that users perform different evaluation techniques and assessments based on the type of website (e.g., e-health or e-government vs. e-commerce).

The fifth and sixth credibility guidelines, ease of contacting the organisation and professional design of the website, support the findings of other studies [2, 3, 4, 8, 9, 10, 25].

The eighth credibility guideline, usability, occupies the highest number of responses: 93 out of 314, or 29.62%.

The frequency of website updates and the presence of advertising material received no responses from the users. Fogg et al. (2003) showed that up-to-date websites represent less than 3% of users responses' regarding credibility. However, the author did not explain whether this result is related to the e-commerce domain or other domains where currencies and up-to-date information are critical, such as in the health sector. Moreover, advertising material plays an important negative role in credibility. In an earlier study, Fogg et al. (2001) mentioned that the well-perceived, reputable ads affect the credibility of the hosted website in a positive way, and poorly perceived ads have a similar negative effect. One possible explanation for this result is that participants are using tools that prevent ad content from appearing. Another possible explanation is that participants performed the study on a secure Internet connection.

Although Fogg's 10 Stanford credibility guidelines have been widely used in the credibility evaluation of websites, they were developed in 2002, which is now more than 15 years ago. Other responses to the open-ended questions included additional criteria not included in Fogg's 10 Stanford credibility guidelines. They are grouped into new credibility criteria guidelines: reputation, security, endorsement and service diversity.

Reputation, one new credibility guideline, refers to all comments related to the brand recognition of the website or products. The comments can be either negative or positive. Surprisingly, reputation occupied sixth place in the most important credibility factors in the e-commerce website evaluations. This finding contradicts the results of previous studies [9], where the reputation of the company was defined as the most important factor in increasing the credibility of e-commerce websites.

The eleventh credibility guideline, security, is a concern that prevents some people from engaging in e-commerce activities. As the customer's sensitive information is transmitted online, security and privacy are key elements to protect the information. Unlike the findings of [9, 10], where no evidence was presented regarding the importance of security, this study confirms that security and information privacy play a vital role in the credibility of e-commerce websites. One possible explanation is that previous research investigated credibility factors in several domains, such as sports websites, news websites, opinion and review websites, and non-profit websites, where no sensitive information needed to be transmitted. In contrary to this study, which focusses on e-commerce websites, information security is of high importance to users.

The endorsement credibility guideline refers to all comments related to what other people think of the website, including customer reviews on the websites, customer reviews on social media, and general recommendations by friends or relatives. The evidence shows that endorsements play an essential role in e-commerce credibility, where people rely heavily on other people's experiences, whether online or offline. This guideline includes both positive and negative comments.

Finally, the service diversity guideline refers to all comments related to the different services a website offers. These services can be identified as services that facilitate the journey of the

customer, from navigating the website to checkout. Such services include shipment estimation, related products and different payment methods.

Fig. 1 represents the frequency of the general credibility guidelines. As noted, the usability guideline obtained 29.62%, making it the most frequent factor that affects the perception of credibility. This is in contrast with the results for English-speaking audiences, for whom professional website design was the most frequent factor that influences users' perception of credibility. As prior studies have indicated, culture is highly important and affects users' perception [6, 15, 37, 38]. The second rank was occupied by information accuracy (20.38%), and professional website design was ranked third, with 14.97%.

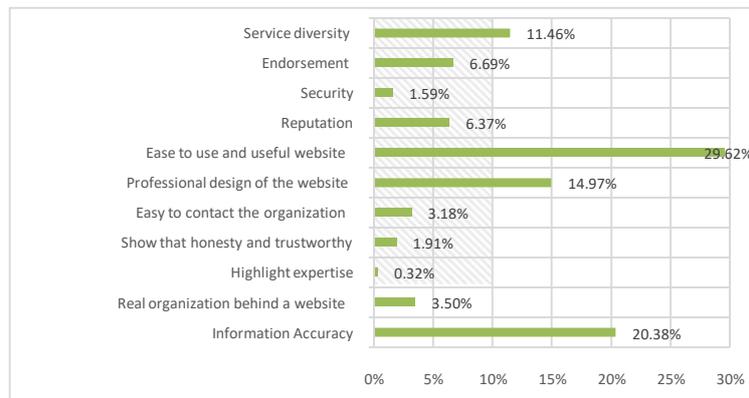


Figure 1. Users' comments regarding credibility

However, the users' first impression of the website credibility was also calculated. The total number of comments in the impression task was 134. As noted, the professional website design obtained 20.15% as the most frequent factor affecting the perception of credibility. Likewise, some previous studies have indicated that aesthetics and professional website design play a vital role in the users' first impression [4, 14]. See Fig. 2.

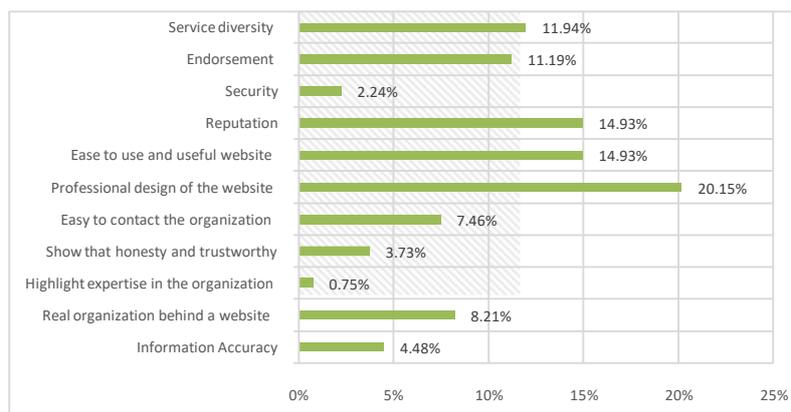


Figure 2. Users' comments in the first-impression evaluation

To answer the first research question, a match of the factors proposed in this study and Fogg’s credibility factors was generated using a confusion matrix Fig. 3. Fogg’s credibility guidelines, with associated criteria from previous studies, are mentioned in Table 1 and are on the X axis in the confusion matrix, while the proposed credibility factors of this study are mentioned in Table 8 and are on the Y axis. As noted, five of Fogg’s factors are not mentioned by the participants in this study: the compatibility between the domain name and the company name, the identification and displaying of any awards, the displaying of company credentials, website updates, and constraints on advertisement material. These criteria are marked red in the confusion matrix.

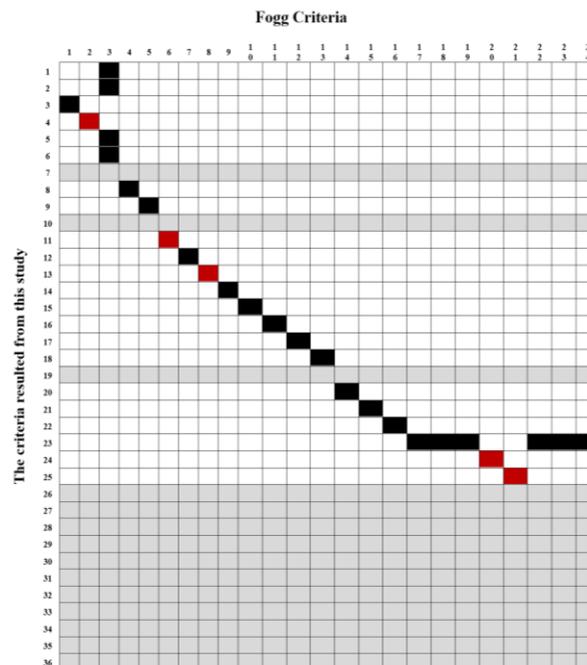


Figure 3. Confusion matrix of Fogg’s guidelines vs. the criteria from this study

This study implemented some additional guidelines: reputation, security, endorsement and service diversity. The reputation guideline refers to the opinions and beliefs of other people about the website (if the reputation is good, then it is credible) and includes criteria relating to both product reputation and website reputation. The security guideline refers to the level of protection against threats for both the website and its customers, and it includes criteria relating to the security of the site, its privacy policy and secure payment methods. The endorsement guideline refers to the act of approving and supporting the credibility of the website, and it includes criteria relating to customer reviews on the site itself, customer reviews on social media, and recommendations from known, trusted individuals. These criteria are marked with a horizontal line in the confusion matrix. The additional guidelines and their associated criteria have also been mentioned in previous studies [35, 36]. However, these studies used self-reporting approach rather than an evaluation approach. Whereas the former is based on the participants' perception of credibility without direct interactions with the system, collected through such methods as interviews, the latter is based on actual interactions between the participants and the system. One major drawback of the self-reported approach is that the users' perceptions while using the site are different from when they are just thinking about it.

Given the factors behind increasing the credibility of local e-commerce websites, this study has confirmed that professional website design and high-quality pictures play a vital role in consumers' first impression of local e-commerce websites. However, usability seems to be the most mentioned factor for consumers when evaluating the credibility of local e-commerce sites. This finding contradicts previous studies [9], where the reputation of the company is defined as the most important factor in raising the credibility of e-commerce websites. This finding is promising, as it encourages and promotes small to medium-sized organisations to boost their businesses from offline to online platforms and pay more attention to website functionality. One possible explanation for this finding might be that Saudi Arabia is a developing country in the field of e-commerce, and the majority of local e-commerce websites have a narrow reputation, with local consumers often seeking local resources [39], which lead the consumers to neglect the website reputation. With the rapid development of communication technologies, it seems that social media play an important role in users' perception of credibility in e-commerce. Criteria such as the website using social media to market itself and the website presenting customers' opinions through social media increase websites' credibility. Furthermore, having dedicated mobile-phone applications also seems to increase websites' credibility.

For the current study, new criteria were produced that were not found in previous studies: using different payment methods, using different shipment companies and using different website services, such as a shipment-tracking service and a related-product service. These new criteria are labelled under the "service diversity" guideline. The absence of these guidelines in previous studies may be due to most of the research in the e-commerce website credibility area having been conducted more than one decade ago, when the services provided by e-commerce websites were highly simple. In contrast, the services provided these days are more intelligent and better understand the behaviour of the consumer. One strong service example is the "customers who bought this also bought this" feature, where collaborative filtering is used to understand the customer's needs and suggest related items. Providing such services facilitates the customer's journey on the website and provide an impression of credibility, as it is impossible to provide such an expensive service on fraudulent websites. As a result, Fogg's 10 credibility guidelines can be implemented in the domain of e-commerce in Saudi Arabia with minor modifications and expansions.

To answer the second research question, according to the one-way ANOVA conducted in study 1, e-commerce1 has the best overall credibility evaluation, followed by e-commerce2, with e-commerce3 scoring the lowest. The results from the one-way ANOVA in study 2 indicated that e-commerce1 has the best overall performance, with the lowest number of clicks and smallest amount of time required to finish the tasks. Users of e-commerce 2 demonstrated the worst overall performance in terms of the number of clicks, and e-commerce3 showed the worst overall performance in terms of the time needed to complete the tasks. Based on these results, we can conclude that user performance is associated with differences in website credibility.

8. CONCLUSION

The results presented in this study shows that Fogg's 10 Stanford credibility guidelines are applicable to Saudi Arabian e-commerce websites. Nonetheless, slight variations exist due to cultural differences, the strong presence of social media, and technological improvements over the years. Additional credibility guidelines (i.e., service diversity, security, reputation and endorsement) were also generated for the current study. The key assertions highlighted in this

study are that professional website design plays an important role in the initial evaluation of the website inspection, while usability is the factor most frequently mentioned by users as affecting their perception of credibility. The second aim of this study was to investigate whether user performance is affected by differences in website credibility. The results indicate an association between users' performance and website credibility differences. The more credible the website, the better users' performance and vice versa. This paper has provided a deeper insight into users' perception of credibility, and understanding users' perception of e-commerce credibility can help in making user interactions with e-commerce more convenient. The results can also help stakeholders to build more credible local e-commerce websites.

REFERENCES

- [1] B. J. Fogg and H. Tseng, "The elements of computer credibility," presented at the Proceedings of the SIGCHI conference on Human Factors in Computing Systems, Pittsburgh, Pennsylvania, USA, 1999.
- [2] P. B. Lowry, D. W. Wilson, and W. L. Haig, "A Picture is Worth a Thousand Words: Source Credibility Theory Applied to Logo and Website Design for Heightened Credibility and Consumer Trust," *International Journal of Human-Computer Interaction*, vol. 30, pp. 63-93, 2014/01/02 2014.
- [3] T. Nind, "Can increasing surface credibility improve e-health intervention effectiveness?," PhD, school of computing university of Dundee 2012.
- [4] D. Robins and J. Holmes, "Aesthetics and credibility in web site design," *Information Processing & Management*, vol. 44, pp. 386-399, 1// 2008.
- [5] N. J. Lightner, "What users want in e-commerce design: effects of age, education and income," *Ergonomics*, vol. 46, pp. 153-168, 2003/01/01 2003.
- [6] Z. Liu and X. Huang, "Evaluating the credibility of scholarly information on the web: A cross cultural study," *International Information & Library Review*, vol. 37, pp. 99-106, 2005/06/01 2005.
- [7] J. Stanford, E. Tauber, B. Fogg, and L. Marable, "Experts vs. Online Consumers: A Comparative Credibility Study of Health and Finance Web Sites," *consumer webwatch2002*.
- [8] F. Alsudani and M. Casey, "The effect of aesthetics on web credibility," presented at the Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology, Cambridge, United Kingdom, 2009.
- [9] B. J. Fogg, C. Soohoo, D. R. Danielson, L. Marable, J. Stanford, and E. R. Tauber, "How do users evaluate the credibility of Web sites?: a study with over 2,500 participants," presented at the Proceedings of the 2003 conference on Designing for user experiences, San Francisco, California, 2003.
- [10] M. Kąkol and R. Nielek, "What affects web credibility perception? An analysis of textual justifications," *Computer Science*, vol. 16, 2015-09-07 2015.
- [11] B. Fogg, J. Marshall, T. Kameda, J. Solomon, A. Rangnekar, J. Boyd, et al., "Web credibility research: a method for online experiments and early study results," presented at the CHI '01 Extended Abstracts on Human Factors in Computing Systems, Seattle, Washington, 2001.

- [12] B. Fogg, J. Marshall, A. Osipovich, C. Varma, O. Laraki, N. Fang, et al., "Elements that affect web credibility: early results from a self-report study," presented at the CHI '00 Extended Abstracts on Human Factors in Computing Systems, The Hague, The Netherlands, 2000.
- [13] S. M. Choi and N. J. Rifon, "Antecedents and Consequences of Web Advertising Credibility," *Journal of Interactive Advertising*, vol. 3, pp. 12-24, 2002/09/01 2002.
- [14] L. A. Peracchio and D. Luna, "The role of thin-slice judgments in consumer psychology," *Journal of Consumer Psychology*, vol. 16, pp. 25-32, 2006.
- [15] R. Khaled, P. Barr, J. Noble, R. Fischer, and R. Biddle, "Our place or mine? Exploration into collectivism-focused persuasive technology design," in *International Conference on Persuasive Technology*, 2006, pp. 72-83.
- [16] B. Hilligoss and S. Y. Rieh, "Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context," *Inf. Process. Manage.*, vol. 44, pp. 1467-1484, 2008.
- [17] B. J. Fogg, *Persuasive Technology: Using Computers to Change what We Think and Do*: Morgan Kaufmann Publishers, 2003.
- [18] A. J. Flanagin and M. J. Metzger, "The perceived credibility of personal Web page information as influenced by the sex of the source," *Computers in Human Behavior*, vol. 19, pp. 683-701, 2003/11/01/ 2003.
- [19] M. J. Metzger, A. J. Flanagin, K. Eyal, D. R. Lemus, and R. M. McCann, "Credibility for the 21st Century: Integrating Perspectives on Source, Message, and Media Credibility in the Contemporary Media Environment," *Annals of the International Communication Association*, vol. 27, pp. 293-335, 2003/01/01 2003.
- [20] Z. Huang, "Usability and credibility evaluation of electronic governments: users' perspective," PhD, School of Information System, Computing and Mathematics, Brunel University, 2010.
- [21] C. a. I. T. C. CITC, "Report E-commerce in Saudi Arabia " 2017.
- [22] N. Clewley, S. Y. Chen, and X. Liu, "Evaluation of the credibility of internet shopping in the UK," *Online Information Review*, vol. 33, pp. 805-826, 2009.
- [23] J. Nielsen. (June 4, 2012, July 3ed). How Many Test Users in a Usability Study? Available: <https://www.nngroup.com/articles/how-many-test-users/>
- [24] B. J. Fogg. (May 2002 21 Feb). Stanford Guideline for Web Credibility. Available: www.webcredibility.org/guideline
- [25] C. N. Wathen and J. Burkell, "Believe it or not: Factors influencing credibility on the Web," *Journal of the American Society for Information Science and Technology*, vol. 53, pp. 134-144, 2002.
- [26] C. WebWatch and P. S. R. Associates, *A Matter of Trust: What Users Want from Web Sites : Results of a National Survey of Internet Users for Consumer WebWatch: Consumers Union of U.S.*, 2002.
- [27] B. J. Fogg, T. Kameda, J. Boyd, J. Marshall, R. Sethi, and M. Sockol, "Stanford-Makovsky Web Credibility Study 2002: Investigating what makes web sites credible today," 2002.

- [28] J. Collins, "An investigation of web-page credibility," *J. Comput. Sci. Coll.*, vol. 21, pp. 16-21, 2006
- [29] K. S. Park and C. Hwan Lim, "A structured methodology for comparative evaluation of user interface designs using usability criteria and measures," *International Journal of Industrial Ergonomics*, vol. 23, pp. 379-389, 1999/03/20/ 1999.
- [30] M. Matera, M. F. Costabile, F. Garzotto, and P. Paolini, "SUE inspection: an effective method for systematic usability evaluation of hypermedia," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 32, pp. 93-103, 2002.
- [31] S. M. Z. Ahmed, C. McKnight, and C. Oppenheim, "A user-centred design and evaluation of IR interfaces," *Journal of Librarianship and Information Science*, vol. 38, pp. 157-172, 2006/09/01 2006.
- [32] J. Nielsen and R. Molich, "Heuristic evaluation of user interfaces," presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Seattle, Washington, USA, 1990
- [33] R. A. Virzi, "Refining the test phase of usability evaluation: how many subjects is enough?," *Hum. Factors*, vol. 34, pp. 457-468, 1992.
- [34] J. S. Dumas and J. C. Redish, *A Practical Guide to Usability Testing*: Intellect Books, 1999.
- [35] E. Bonsón Ponte, E. Carvajal-Trujillo, and T. Escobar-Rodríguez, "Influence of trust and perceived value on the intention to purchase travel online: Integrating the effects of assurance on trust antecedents," *Tourism Management*, vol. 47, pp. 286-302, 4// 2015.
- [36] A. Elwalda, K. Lü, and M. Ali, "Perceived derived attributes of online customer reviews," *Computers in Human Behavior*, vol. 56, pp. 306-319, 2016/03/01/ 2016.
- [37] K. Kang, "Considering Culture in Designing Web Based E-commerce," in *E-commerce*, ed: InTech 2010, p. 15.
- [38] P. Y. K. Chau, M. Cole, A. P. Massey, M. Montoya-Weiss, and R. M. O'Keefe, "Cultural differences in the online behavior of consumers," *Commun. ACM*, vol. 45, pp. 138-143, 2002.
- [39] Mastercard. (2014, July 25). *Online Shopping Behavior in Saudi Arabia*. Available: <https://newsroom.mastercard.com/mea/press-releases/online-shopping-on-the-rise-in-saudi-arabia-mastercard-survey/>
- [40] S. Sait, K. Al-Tawil, and S. Hussain, "E-commerce in Saudi Arabia: Adoption and perspectives," *Australasian Journal of Information Systems*, vol. 12, 2004
- [41] R. AlGhamdi, O. A. Alfarraj, and A. A. Bahaddad, "How Retailers at different Stages of E-Commerce Maturity Evaluate Their Entry to E-Commerce Activities?," *arXiv preprint arXiv:1503.05172*, 2015
- [42] PayFort, "State of Payments 2016," Report 2016.
- [43] B. Fogg, J. Marshall, T. Kameda, J. Solomon, A. Rangnekar, J. Boyd, *et al.*, "Web credibility research: a method for online experiments and early study results," presented at the CHI '01 Extended Abstracts on Human Factors in Computing Systems, Seattle, Washington, 2001.