

Investigation into the Impact of the Usability Factor on the Acceptance of Mobile Transactions: *Empirical Study in Saudi Arabia*

Mohammed A. Alqahtani

*School of Computing Sciences
University of East Anglia
Norwich, NR4 7TJ UK
College of Computer Science and Information Technology
University of Dammam
Saudi Arabia*

m.alqahtani@uea.ac.uk

Obead Alhadreti

*School of Computing Sciences
University of East Anglia
Norwich, NR4 7TJ UK
College of Computer
Umm Al-Qura University*

O.Alhadreti@uea.ac.uk

Roobaea S. AlRoobaea

*School of Computing Sciences
University of East Anglia
Norwich, NR4 7TJ UK
Faculty of Computing and Information Systems
Taif University
Saudi Arabia*

r.alroobaea@uea.ac.uk

Pam J. Mayhew

*School of Computing Sciences
University of East Anglia
Norwich, NR4 7TJ UK*

p.mayhew@uea.ac.uk

Abstract

Both the rapid advances in wireless technologies and the high proliferation rate of mobile communication services and artefacts have had a profound impact on industry, and are beginning to offer interesting and advantageous new services. In particular, the mobile transaction (m-transaction) system has emerged, enabling users to pay for physical and digital goods and services using their mobile devices whenever they want, regardless of their location. Although it is anticipated that m-transactions will enjoy a bright future, there is apparently still reluctance among users to accept mobile transactions, particularly in Saudi Arabia. Furthermore, it is believed that usability is critical to the success and acceptance of mobile technology innovations in general, and m-transactions in particular. While most of the research into IT innovation acceptance is based on TAM (Technology Acceptance Model) and its many variants, the literature review reveals that there is little research on mobile application usability as an important antecedent factor for the success and acceptance of mobile technologies. This is particularly true in developing countries, especially the Middle East, therefore this project aims to fill this gap in the research.

This paper is an extension of four previous studies by the authors, which indicated that usability is the most influential factor in Saudi users' intentions of utilising mobile transactions. It aims to investigate and empirically test the usability level of mobile transactions from the perspective of Saudi users; this will be achieved by conducting usability experiments with real mobile

applications within the target group. In other words, this study is a part of a larger research project and it serves as a validation/verification tool to answer the question "Why does the ease of use of mobile transactions have the highest impact on the intention to use mobile transactions in Saudi Arabia?" Moreover, it will produce a set of empirically-based recommendations for the developers and providers of mobile applications who are interested in the Saudi market. The results of this study showed that the usability of the tested application, which is considered one of the most popular commercial mobile application in Saudi Arabia, (i.e. 'Souq.com'), is unsatisfactory, with an average SUS Score of 50. Not surprisingly, this low usability negatively affected the willingness of the participants to use the application again.

Keywords: Mobile Transaction, Mobile Usability, Souq.com, IT Innovation Acceptance, Human Computer Interaction.

1. INTRODUCTION

Since the invention of the telephone in the 1870s, the remarkable advancement of wireless technology has profoundly changed the telephony system. Mobile devices now have functions that surpass mere telephony needs, and which motivate the development of value-added mobile services and functions such as: social networking, mobile banking, entertainment, health and fitness applications. Mobile commerce (m-commerce) involves the trade of goods, services and content through mobile devices, without any limitations of time or place – known as 'anytime-anywhere'. Furthermore, Cheong and Park [1] state that "m-commerce is electronic commerce over wireless devices and requires a transaction of monetary value through a wireless telecommunications network. Without transactions of monetary value, m-commerce cannot be realised". As m-commerce technologies evolve and their popularity increases, mobile payments should be able to facilitate the provision of secure electronic transactions between organizations and individuals. The number of mobile devices being used far exceeds any of the other technical devices that can be used to market, sell, produce, or deliver goods and services to consumers. This development provides very lucrative opportunities to merchants and service providers [2]. As a result, mobile communication technologies have successfully penetrated consumer markets throughout the world. Therefore, mobile transactions play an essential role in mobile commerce, and they are a crucial factor in the success of mobile commerce penetration [3, 4].

The researchers extensively reviewed the existing literature; this revealed that there are no studies investigating the usability of mobile applications in Saudi Arabia, or how the usability could affect users' intentions to employ m-transactions. That makes this study a first in its field. This study commenced in April 2011 and is a part of a larger research project into m-transaction acceptance by users in Saudi Arabia. The first phase was an exploratory study of e-commerce (the first wave of m-commerce) and was conducted in OSaudi Arabia. The main aim of this exploratory study was to investigate, from the consumer's perspective, the main factors that play a role in the adoption of e-commerce. A 'grounded theory' methodology was used to collect and analyse the data. Semi-structured interviews were conducted with Saudi residents to elicit their opinions about e-commerce enablers and disablers in Saudi Arabia. The findings of the study suggest that the most influential factors on the adoption of e-commerce in Saudi Arabia are: security, fraud and hacking, trust, cyber-laws, awareness and perceived usefulness, postal services, government e-readiness, resistance to change, the presence of commercial electronic websites, cost, tangibility, warranty, trial and experience. The study was published in 2012 [5]. The second phase of this project focused on m-transactions; its overall aim was to investigate the key factors that affect the adoption and intention of using m-transactions from the consumer's perspective. From this, the researchers were able to enhance and improve the developed framework which encompasses and categorises the influential factors of m-transactions. The findings of this study show that the most influential factors on the acceptance and use of m-transactions in Saudi Arabia are: usability, usefulness, telecommunications infrastructure, security, hacking and fraud, availability, trust, payment gateway, awareness, cost and promotion, privacy, cyber-laws, the postal services, government e-readiness and Arabic language support. This study

was published in 2014 [4]. In phase 3, by analysing and discussing the author's prior research (i.e. the first two phases), this study aims to identify the factors that influence the acceptance of mobile transactions (m-transactions), and to build the conceptual framework of the intention to use m-transactions in Saudi Arabia. It also summarizes the factors that should have a significant impact on the intention to use m-transactions from a consumer's perspective. The results show that there are 11 factors that have the most significant influence on the acceptance and adoption of m-transactions in Saudi Arabia: ease of use, navigational structure, visual appeal, usefulness, ICT infrastructure, security, trust, culture, cost, government m-readiness, and social influence. This study was also published in 2014 [6]. In phase 4, the developed framework was empirically validated in a study which involved a statistically representative sample size of approximately 1,000 Saudi users from different demographic backgrounds. The empirical analysis revealed that security, ease of use, individualism, masculinity, navigational structure, power distance (strength of social/employment hierarchy), uncertainty avoidance, usefulness, and website support for mobiles all have a significant impact on consumers' intentions to use m-transactions. Amongst these factors, ease of use was the most influential. This study has been started in 2014 and the authors aim to publish its results as soon as possible.

The research results showed that the usability of m-transactions was the most important concern for Saudi users [4, 5, 6]. In particular this study aimed to validate this result by conducting a usability test for m-transactions with Saudi users. The feedback from these tests should help to elucidate the results and shed some light on why usability was ranked higher than other factors.

This paper is organised as follows: Section 2 provides a literature review of mobile e-commerce adoption and usability in several countries; Section 3 describes the methodology base from which the findings can be problematized and put into context; Section 4 describes the research measurements and design; Section 5 presents the results and discussion; Section 6 offers a conclusion of the significant results, illustrating the factors that affect the decision to use m-transactions and outlining future work; and the final section discusses the novelty of this research.

2. LITERATURE REVIEW

Despite the growing interest in m-commerce platforms, the poor usability of mobile sites and applications for commercial activities is a prominent major obstacle to the adoption of mobile solutions. "Such difficulty discourages users from accessing mobile Internet sites," [7]. Furthermore, Min and Li [8] highlighted the importance of usability as a factor for adopting mobile commerce; their study in China confirmed that usability is crucial to consumers' use of mobile transactions. Buranatrived and Vickers [9] revealed that usability acts as a second barrier, after security, to the acceptance of m-commerce. Venkatesh, Ramesh and Massey [10] in the USA also investigated the usability of mobile sites, identifying it as a significant requirement for successful m-commerce applications. Furthermore, they emphasised that the usability of Internet-based commercial sites on the traditional Personal Computer (PC) should not necessarily be equated directly with m-commerce and thus, special attention should be paid to m-commerce usability. Their study, based on surveys of mobile Internet users, concluded that usability is the biggest source of frustration for consumers. Thus, a vital requirement for the success of m-commerce applications is first understanding which aspects of usability are important to users, and how they may differ in a wireless context; this will help to ensure that the customers' experiences are positive. Another study by Li and Yeh [11], conducted in Taiwan across two large cities in three universities, analysed a total of 200 responses from distributed surveys to gain an understanding of the factors affecting users' trust in mobile devices. Their research discovered that design aesthetics had a significant impact on perceived usefulness and ease of use, and that all of these were ultimately shown to have an important effect on customers' trust in mobile commerce [11]. [12] pointed out that mobile devices create major challenges for investigating the usability of mobile applications, due to factors such as their small screen size. Nielsen and Budiu [13] in their book "Mobile Usability" referred to the fact that the percentage of visiting users who successfully accomplish their desired action (conversion rate) is very low in mobile users (1.4%) compared to users of personal computers (PCs) (3.5%). They offered two reasons for this dramatic difference: the mobile users' experiences (usability) often lead to frustration, which

results in the experiences being ranked below the users' expectations; in addition, some companies perceive mobile site design as an area not worth investing in. Nielsen and Budiu recommended that companies should optimise their mobile sites and enhance their usability, thus increasing the conversion rate of mobile users to a competitive level of that of PC users [13]. The system/product experience varies from person to person due to the concept of 'usability'. Usability is a core term in Human-Computer Interaction (HCI) and it can be defined as "Ease of use and acceptability of a system for a particular class of users carrying out specific tasks in a specific environment" [14].

There are many usability evaluation methods (UEMs) which can be employed to discover usability problems in a system, and to measure the success of users in achieving particular goals within a specified context. In-lab usability testing has been used as the standard method for evaluating and improving the usability of mobile devices. Indeed, a meta-analysis study revealed that 71% of mobile device evaluations were conducted in lab-based settings [15]. With this method of testing, a mobile application is evaluated in a prepared environment. The testing sessions are recorded by cameras, and the recordings are analysed by the evaluators [16]. Think-aloud (TA) protocol is commonly employed as a complementary technique with in-lab testing. Users are encouraged to verbalise their thoughts while performing a predefined set of tasks [17]. They are carefully observed by evaluators, who analyse their behaviour and record the problems encountered by them; these are then compiled into a list of usability problems. Metrics such as the number of errors, the total time spent, the success rate and user satisfaction are also recorded for further analysis. The usability problems uncovered by this process can then be corrected to improve the quality of the product [18]. Furthermore, there are three TA approaches: concurrent, retrospective and constructive interaction. The concurrent approach is the most commonly-used for usability testing; it involves participants verbalising their thoughts whilst performing tasks in order to evaluate a system. The retrospective approach is less frequently used; it involves participants accomplishing their tasks silently, and then afterwards commenting on their work, prompted by a recording of their performance. After the experiment, experts can review all the recordings to uncover the usability problems. Constructive interaction is more generally known as Co-Discovery Learning, where two participants work together to perform a task, verbalising their thoughts while interacting [19].

Field studies are another type of UEM, which in contrast to lab-based testing, take into account the dynamic mobile context and unreliable wireless networks. This method involves recording observations and interviewing users, with the aim of understanding their requirements of the product [20]. However, it has some limitations, such as problems that can arise with the use of video recordings or observations, difficulties in gathering data, and the inability to control conditions; the researchers need to physically move around in a dynamically changing environment. In addition, it may be problematical to create realistic environments that capture the richness of the mobile context [12].

The ultimate results of applying either user testing (UT) or TA approaches is a list of usability problems, which then need to be categorised in some way to enable further analysis. The problems can be classified into different groups, using a numeric scale to measure the severity of each problem. For example, Nielsen developed a rating scale which consists of five scales: '0' means that this issue is not a usability problem at all; '1' signifies that this is a cosmetic problem which does not need to be fixed, unless extra time is available on the project; '2' means that this issue is a minor usability problem, and fixing it should be given low priority; '3' corresponds to a major usability problem which should be given high priority, and finally '4' signifies that this issue is a usability catastrophe, and that it is imperative to fix it before the product is released [21].

The researchers of this paper conducted meta-analysis in which they looked into the different studies carried out on m-transaction usability and adoption, as summarised in Table 1.

Literature	Research Purpose	Methods Used	Theory/Constructs Used	Results/Finding
[22]	To investigate mobile usability in mobile “private shopping” applications.	A qualitative mobile usability test, based on a multi-method approach with a sample of 11 Turkish senior year university students.	Think aloud, eye-tracking and video recording.	Significant mobile usability problems.
[8]	To explore the adoption of m-commerce from the perspective of system usability.	Extensively review the literature.	Mobile device characteristic; WAP web characteristic; wireless communication characteristic; context; m-commerce system usability.	A comprehensive framework to help researchers understand usability issues systematically in m-commerce.
[23]	To identify the determinants of pre-adoption of mobile payment services and explore the temporal evolution of these determinants between the pre-adoption and post-adoption stages from a holistic perspective.	Online survey 639 Respondents (483 potential adopters + 156 adopters).	Social influence; personal innovativeness; perceived risk; perceived fee; compatibility; relative advantage → behavioural intention to adopt.	Behavioural beliefs in combination with social influences and personal traits are all important determinants for mobile payment services adoption and use. However, their impact on behavioural intention varies through different stages.
[24]	To examine the factors influencing the intention to use mobile payment.	Distributed survey + email survey. 269 respondents who have experience of mobile payments.	Extended Technology Acceptance Model (TAM): perceived usefulness; perceived ease of use; two consumer-centric factors (i.e. <i>personal innovativeness</i> and <i>m-payment knowledge</i>) and four m-payment system characteristics (i.e. <i>mobility, reachability, compatibility, and convenience</i>) → intention to use m-payment.	The strongest predictors of the intention to use m-payment are ‘perceived ease of use’ and ‘perceived usefulness’.
[25]	To explore the factors determining consumers’ acceptance of mobile payment services.	Online survey 1447 responses (583 with experience of mobile payment services, and 864 with no experience).	Extended Technology Acceptance Model (TAM): perceived compatibility; perceived security; perceived usefulness; perceived ease of use; individual mobility; subjective norm; attitude towards use → intention to use.	Perceived compatibility has the greatest impact on the intention to use mobile payment services; individual mobility and subjective norm also have a significant influence.
[26]	To develop a model which explores how customers’	Distributed questionnaires. 263 respondents	Service (convenience, functionality, affordability and	Cost and individual's economic status are important decision-

	perceptions of the value offered by m-payment services influence their attitudes towards m-payment adoption.	(46.1% male and 53.56% female + one did not specify the gender).	service awareness); technology (interface, self-efficacy and security); perceived usefulness; perceived ease of use → intention to use → use.	making factors. The familiarity and general awareness of the mobile payment service affect positively the intention to use it.
[3]	To investigate the consumers' acceptance of mobile payments in KSA.	Distributed survey 200 respondents who are mobile phone users.	Security; unauthorised use of phone to make purchase; complexity; and limited amount of money per transaction.	Security of m-payment and unauthorised use of mobile phones are the strongest concerns.
[27]	To review prior literature on mobile commerce, analyse the various adoptive factors, and suggest future research direction.	Conducting an exhaustive and systematic electronic search of 59 articles (23 Chinese and 36 English).	Research subjects; research theory; and adoptive factors.	Most of the existing research is on consumer adoption; TAM is the most-used theory in consumer adoption; the majority of adoptive factors are characteristics of mobile commerce service providers, merchants, and consumers; the environment; and consumers' perceptions of features of mobile commerce

TABLE 1: Summary of Different Studies on M-Transaction Usability and Adoption.

From the above summary, it is clear that the usability of m-commerce is a significant factor considered by consumers when they are deciding whether or not to adopt m-transactions. This has been investigated in different countries, for example USA, China, Turkey and Taiwan, however no research has so far confirmed this from the perspective of Saudi users, which is the issue that this research will address.

3. METHODOLOGY

3.1 Evaluation Methods

Choosing the correct evaluation method is important; scientifically validated information on appropriate testing methods is valuable for usability practitioners. The Thinking-Aloud protocol is one of the most important usability evaluation methods (UEMs). It has been a key evaluation framework ever since the 1980s, and it remains important today in the Information System (IS) field [28, 29, 30]. In this study, user testing and think-aloud 'concurrent approach' methods were chosen. These methods are the best way to examine whether usability is the critical factor that affects the intention of using m-commerce among Saudi people.

3.2 Test Application

Firstly, the researchers sought to ensure that the selected mobile application would support the research goals and objectives. The selection process was criteria-based, and five aspects were determined and verified: (1) fully working application with rich functionality; (2) the whole processes can be conducted in a mobile context (e.g. via smart mobile phone); (3) the application was developed and intended for use in the Saudi market and by Saudi users; and (4) it has not been used before by the test participants. The researchers faced some difficulties when trying to

find an application which met all four criteria and which would provide a full transaction via the mobile application – i.e. (1) registering with the application; (2) shopping and finding the required product; (3) adding the product to the basket and managing the basket; (4) conducting the purchase transaction; (5) contacting the support team about the purchased product. The researchers could only find one mobile application which was available in the Saudi market and which met all four criteria: the well-known application called ‘Souq.com’. Souq.com is considered the largest e-commerce site in the Arab world, featuring more than 400,000 products across many different categories (e.g. fashion, electronics, household goods ...etc) and is often tagged as the “Amazon of the Middle East”. It was established in 2005 as an auction site linked to the Arab internet portal “Maktoob” then moved to a marketplace with a fixed price model in early 2011, and finally launched its retail division at the end of 2011 [31]. The researchers were able to gain approval by email for conducting the usability study on the mobile application for Souq.com, and it was decided to conduct the test on an IOS device, as it is one of the most popular mobile operating systems [16].



FIGURE 1: The tested mobile application “Souq.com” [31].

3.3 Tasks

In-lab usability testing requires representative users of the product to be recruited to perform a set of predesigned tasks. These tasks should be as representative as possible of the activities that real users would perform on the selected application. The tasks should be designed to focus on the main functions of the system, such as the product and catalogue pages, searching features, displaying of records, interactivity, participation and sorting features [32]. Snyder [33] outlined the characteristics of tasks which would be suitable for use in testing, which can be summarised into 6 points: (1) based on a goal that matters to the user’s profile; (2) covers important questions about the success of the product; (3) has appropriate scope; (4) has predictable and a finite number of possible solutions; (5) has a clear end that the user can recognise; and (6) elicits action. Furthermore, Snyder [33] proposed a task template which can be used for any type of usability testing, shown in Table 2. This template was carefully applied to all the experiment’s tasks.

Task #: < Task Name >
Goals/output:
Inputs/Assumptions:
Steps:
Time for expert:
Instructions for user:
Notes:

TABLE 2: Task Template [33].

Furthermore, a usability context of use analysis is commonly used before designing the tasks, in order to ensure that the tasks are realistic and that the users are representative of a typical user. This analysis usually mitigates the artificial nature of usability testing and improves the validity of the results. The usability context of use analysis can be carried out in various ways. A proper way to collect the information is to arrange a meeting with stakeholders who have knowledge about the system under evaluation. If it is impossible to arrange a meeting, the information can be collected by interviewing the stakeholders or by using a questionnaire. Stakeholders may include, but are not limited to, the project manager, users' representatives, developers and system support [34]. In the absence of any response from the "Souq.com" owners/representatives regarding the main functions of their application, the researchers arranged interviews with five active "Souq.com" members. An interview agenda containing the topic of the study, the aim of the interview, and the estimated time was sent to the interviewees beforehand. The interview had a default time limit of 10 to 15 minutes in order to maintain an active conversation. The interviews took place in a convenient place predetermined by the interviewee (e.g. in a coffee shop near to the interviewee's house) and involved two main questions: (i) what type of tasks do you most commonly perform in Souq.com that involve mobile transactions? (ii) If you have any further ideas/suggestions on any relevant topic, please feel free to let us know. The researchers used the information acquired in the interviews to create seven different test tasks. Furthermore, two experts were consulted to evaluate the tasks. One task (downloading the application from the Apple store) was removed based on the experts' recommendations, as this task involved evaluating the store holding the application – e.g. the Apple store itself – rather than the actual application.

There are a number of different categories of usability testing tasks; the best known are: structured tasks, uncertain tasks, and scenario tasks (also known as problem solving). The tasks adopted in this study were developed and presented in a scenario format. Scenario task formats are the most widely used in usability testing, and are usually recommended by usability textbooks such as [35, 36], as they help usability testing to emulate real-world contexts of use. For further details of the tasks, such as their titles, instructions, goals, input, steps, times and notes, please see *APPENDIX A*.

The task time is measured by recording the time it takes a user to perform a predefined task; this can be the total time spent on a web page or the duration of a phone call. It could be measured in seconds, minutes, hours, days, months or even years [37]. In order to evaluate the usability of a task, the researchers needed to set a maximum acceptable task completion time (specification limit), against which the actual time for completion of the task could be compared. Setting this limit is not easy. Guidelines in the usability literature suggested four approaches for setting task time specification limits [38]:

- 1- The task is examined by the test designers and they set the criteria.
- 2- The fastest task time is identified (can be by an expert), and then the specification limit is set to 1.5 times this time for each task.
- 3- Historical tests with the system can be used to set the time limit.

- 4- An agreement for the time specifications can be reached, based on negotiations between all parties responsible for the system.

Therefore, the researchers decided to set the acceptable time limit for the 6 tasks by applying two methods:

- 1- The tasks are completed by two experts and the formula (expert 1 time * 1.5) is used.
- 2- The tasks are completed by two normal users; their times are added to the times of the two experts, and the mean is calculated.
- 3- A final decision was taken by the experts based on a discussion for each task.

Table 3 shows the time of each task for the two experts (Expert 1 & 2), the tasks time for the two users (User 1, User 2), the average time for the experts and users, the time for Experts 1 & 2 times 1.5, and the final agreed tasks times.

	Expert 1	Expert 2	User 1	User 2	Time average	Expert 1 * 1.5	Expert 2 * 1.5	Tasks Times
Task 1	1.19	1.55	1.40	2.30	1.61	1.79	2.33	02:01
Task 2	1.30	1.28	1.32	1.30	1.30	1.95	1.92	01:30
Task 3	0.36	0.30	2.25	1.20	1.03	0.54	0.45	01:00
Task 4	0.57	0.45	0.57	0.27	0.47	0.86	0.68	00:46
Task 5	2.30	2.30	4.40	2.30	2.83	3.45	3.45	03:20
Task 6	2.20	2.20	4.19	2.21	2.70	3.30	3.30	03:10

TABLE 3: Tasks Timing.

3.4 Participants

The number and background of test participants are key factors in usability testing. To date, there is no agreement on how many users should be involved in usability testing. Some researchers state that eight participants are sufficient, while others suggest that ten participants are enough for an effective test [39]. A recent study [40] investigated the sample size required for usability testing through empirical studies, with the aim of answering this challenging question that has been hotly debated by researchers for many years. Their results concluded the following: 5 users are enough to uncover cosmetic problems and problems relating to structure and content; 8 users are enough for a few major, and some minor, problems, and are more appropriate for commercial studies and problems with layout and formatting; 16 ± 4 users are enough to discover catastrophic, major, minor and cosmetic problems, and also for finding problems relating to design, navigation and the key aims and functions of the system, particularly in comparative studies. Finally, > 20 users are needed for statistically significant studies and analysis of performance metrics, such as the success rate.

Accordingly, it was decided that 30 participants would be recruited for this study (see Table 4 on following page). In usability testing, the test sample should be as representative as possible of the target users. Relevant users are more likely to encounter relevant problems, which in turn will produce more relevant results. The test sample therefore had to be mobile phone users, be familiar with m-commerce, but without prior experience of the 'Souq.com' app. In addition, two independent experts were recruited to rate each problem and to help with coding and categorizing them.

Demographic	Category	Frequency	Percent
Age	Less than 20	8	26.7
	20 - 34	18	60.0
	35 - 49	3	10.0
	50 - 65	1	3.3
	Total	30	100.0
Income *(SAR)	Less than 50,000	9	30.0
	50,000 – 100,000	5	16.7
	100,001 – 150,000	10	33.3
	150,001 – 200,000	3	10.0
	200,001 – 300,000	3	10.0
	Total	30	100.0
Qualification Level	Under Bachelor	9	30.0
	Bachelor	9	30.0
	Higher Diploma	1	3.3
	Master	8	26.7
	PhD/Doctor	3	10.0
	Total	30	100.0
Mobile Skills level	Intermediate	14	46.7
	Advanced	16	53.3
	Total	30	100.0
How long have you been using mobile internet?	1 - 2 Years	1	3.3
	3- 5 Years	18	60.0
	More than 5 years	11	36.7
	Total	30	100.0
How often do you use mobile transactions?	Every day	2	6.7
	Weekly	11	36.7
	Monthly	10	33.3
	Hardly ever	7	23.3
	Total	30	100.0

TABLE 4: Summary of Respondents' Demographic Data.

*SAR = Saudi Arabian Riyal, 1.00 SAR= 0.266 US.

3.5 Ethical Concerns

This study has passed the UEA Computing Sciences Research Ethics Committee. The participants' information was kept anonymous. The participants were told at the beginning that they were totally free to stop or withdraw from the experiment at any time, without any questions. They could refuse to answer any questions and permission for recording the experiment was gained from the participants prior to starting the tests.

3.6 Materials and Equipment Used

In the experiment preparation phase, a set of preparation documents was built consisting of an information and task sheet, an observer sheet, and a demographics and satisfaction questionnaire, see APPENDIX A, B and E. These documents were handed to the participants prior to the experiment, with the exception of the satisfaction questionnaire which was completed once the user had finished the test tasks. The information sheet contains the goals and objectives of the assessment and the roles of the users. The task sheets contain information about the 6 tasks, organised and divided into seven sections: the first section explains the purpose of each task and how to perform it, and also explains how to transfer from one task to another. The other six sections explain the task goal, and provide instructions and details which the users need to perform the relevant task. The test environment was a quiet room. The researchers tried to identify which equipment the users regularly used and set it up for them before the test. A mobile usability testing sled was developed, as there is a lack of these tools in the market and they are expensive. This gadget provides a perfect angle for the camera to capture the user's behaviour and actions while performing the activities on the smart phone. It also provides a stand for the smart phone to rest on in a comfortable way for the user. A separate microphone with a stand was also set up to give a clearer sound recording, and to make it more

convenient for the user as he/she did not then need to lean towards the camera's microphone. These tools were all attached to a laptop which recorded a video of each task for each user. Microsoft Windows Movie Maker was installed and utilised to import the recording from the camera and the audio from the attached microphone to provide a single multimedia file (video), which could then be edited. The researchers could also watch the videos, and fast-forward, rewind and increase/decrease the speed of each video footage (See Figure 2).

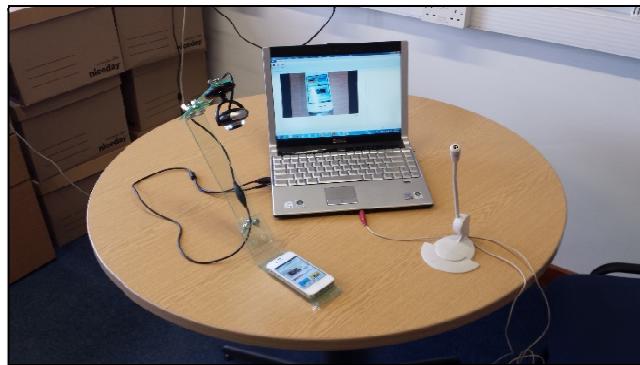


FIGURE 2: Equipment used.

3.7 Piloting and Correction

To make sure that the tasks and all the procedures were clear and understandable, a pilot study was conducted with 4 users. All materials were checked to make sure that there were no spelling or grammatical errors and no ambiguous words or phrases, the users' behaviour was monitored, and all the usability measures were assessed as they would be in the actual testing. All of these steps resulted in useful corrections and adjustments for the real test. Further changes were made based on the users' comments such as rewording some statements and adding some pictures for further illustration.

3.8 Experimental Procedure

The experiments took place in a typical usability lab setting in Saudi Arabia from July 2014 to September 2014. All the participants were selected at random regardless of their age, income and gender. The researchers approached people from private and public organizations in different locations (e.g. universities, companies, schools, Internet cafes, streets and shops). However, this research is about adopting m-transactions, so the researchers had to verify that the participants were mobile users, familiar with mobile commerce, and had conducted a mobile transaction before, but had not used the tested application before (this was verified in the demographic section of the questionnaire [see APPENDIX B]). The sample size of this experiment was 30. One of the researchers (the main author) adopted the role of the evaluator in all the sessions, noting all the comments made by the users. The evaluator researcher used a stopwatch to record the time spent by each user on each task, and the observation sheet to write down the behaviour of each user and the problems discovered. After the users had finished their tasks, they were asked to rate their level of satisfaction in a SUS questionnaire with a five point scale – 1 for strongly disagree and 5 for strongly agree. Bangor, Kortum and Miller [41] suggested slightly modified statements from those used in the original SUS instrument by [42]. The changes are (1) replacing the word "cumbersome" with "awkward" and (2) replacing the word "system" with "product". These studies applied this slightly modified SUS and also replaced the word "product" with "application" as can be seen in Appendix C. To add a valid assessment of what the absolute numerical score of SUS means, a 7 point adjective rating scale can be used to judge the SUS scores. This has become known as the "university grade analog" [41]. This study applied this technique and again replaced the word "product" with "application" (see Appendix C). Each participant conducted the experiment individually and the researcher ensured that every participant felt at ease whilst performing the tasks. Furthermore, users were given a chance to fill in an open-ended questionnaire by writing down their comments and feedback on the application,

and explaining any reaction that was observed during the test. Subsequently, the researcher extracted the problems from the observer sheet and they were merged into a final master problem list.

Before starting the actual evaluation, users completed a demographics questionnaire (see APPENDIX B) to provide the researchers with more information about themselves. Paper-based instructions were given to the users. The moderator described the test environment and how to use the equipment, gave a quick introduction to the task designs, and explained how to 'think aloud' for each test. In addition to the voice recording, video footage of the mobile's screen was captured. All these records were supplemented by the evaluators' notes. The experiment consisted of three stages as follows:

Introductory stage: Before the experiment, participants were introduced to the procedure by instructions read aloud by the evaluators. They were then given a questionnaire to complete with their personal details, demographics, and their experience of using a mobile phone, specifically their experience with mobile transactions. When all the necessary documents had been completed, the participants were allowed ten minutes to explore the Souq.com application. After the allotted time or when the participant said that he/she was ready, the experiment proceeded to the next stage. There are several reasons for allowing exploration of the application: firstly, it helps to familiarise participants with the interface; secondly, it gives the evaluators time to adjust the video recording equipment and test other equipment; finally, the exploration phase serves as a kind of rehearsal before the real experiment that follows. It gives invigilators time to take their seats and prepare for taking notes, while users on the other hand can become used to the hardware and software, and regain their normal speed of interaction with the mobile systems.

Main stage: The main stage of the experiment began with an explanation of the tasks. Each user had a file containing the task sheets, but the tasks were also explained verbally by the evaluator, first by reading the text of the task and then by describing what kind of result he/she expected to see. After this explanation, the evaluator asked the users to explain each task back to him/her using their own words, to check their understanding. The execution of the tasks started after the video recording equipment was activated. The participants performed the tasks at their own pace and executed one after the other. During the entire experiment the evaluators were present without interfering with the tasks. If a participant became silent for a period of time, the evaluator reminded him/her about thinking aloud by saying, "Please think aloud". The users could ask questions before and after execution of the task, but not while the task was being performed.

Post experiment stage: After every experiment, the user was assured of the importance of their participation in the experiment and thanked. After that all the documents and notes relating to that experiment were stapled together and placed in a file. All the video footage of that participant and their screen were compressed and copied to a folder identified by the number of the experiment. Finally, the testing environment was restored back to its original state so that the next experiment could take place.

4. MEASUREMENTS

To determine how the target application had impacted on users' ability to complete their m-transaction, we used the following measures:

- 1- Number and types of usability problems.
- 2- Time spent: this was a measure of the time spent by each user on completing the relevant task. In other words, it was used to measure the efficiency attribute of the application under evaluation.
- 3- Users Satisfaction System Usability Scale (SUS) was used to show how pleased people were with their experience of using the tested application.
- 4- Success rate: this metric is used to measure the users' ability to successfully complete tasks.

5. RESULTS AND DISCUSSION

In this section the analysis and results obtained from the experiments will be presented. This includes a detailed analysis of the problems discovered and their type, participants' performance of tasks, and their satisfaction level.

5.1. Task Performance

5.1.1. Task Completion

As mentioned earlier there were 30 users who conducted the 6 test tasks. Each participant was asked to perform 6 tasks with the targeted application Souq.com, meaning that a total of 180 tasks were performed. At the end of each task, the researcher assessed its completion rate and then classified it as successful (completed) or unsuccessful (not completed). For each task the user was left alone until he/she completed the task, or decided that they could not perform the task (they gave up). All the tasks were either completed successfully or not completed; there were no partially successful tasks. Table 5 below shows the completion rate for each task in numbers and in percentages. Furthermore, it shows those who successfully performed the task within the time allocated to each task (benchmark) and those who successfully performed the task but exceeded the time limit for that task (without benchmark).

	Participants performing correctly (with benchmark)					Participants performing correctly (without benchmark)	
	# Users	#	%	Time Average	SD	#	%
Task 1	30	9	30	03:02	0.001	13	43.33
Task 2	30	3	10	02:29	0.001	15	50
Task 3	30	3	10	03:19	0.001	4	13.33
Task 4	30	13	43.33	01:15	0.001	10	33.33
Task 5	30	2	6.67	06:09	0.002	3	10
Task 6	30	3	10	03:51	0.001	2	6.67

TABLE 5: Task Completion vs Benchmark.

It can be seen from the above table that in Task 1 there were 9 participants (30%) who were able to successfully complete the task within the time frame, however 13 users (43.33%) needed more time to finish the allocated task. The table also shows that the number of participants completing the task within the benchmark time were 3, 3, 13, 2 and 3 for Task 2, Task 3, Task 4, Task 5 and Task 6 respectively. On the other hand, 50% of the users were unable to complete Task 2 inside the allotted time, as were 4, 10, 3 and 2 users for Tasks 3, 4, 5 and 6 respectively. Also, from the average time taken for each task as shown in the table above, we can see that Tasks 5 and 6 took the longest average time, while Task 4 had the shortest average time. The standard deviation between users for each task is low which indicates that there is no significant fluctuation between the users' performances, confirming that the test environment was stable.

	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Total	Average
Successful	22	18	7	23	5	5	80	13.33
Unsuccessful	8	12	23	7	25	25	100	16.67
Total	30	30	30	30	30	30	180	30.00

TABLE 6: Task Completion.

The above table (Table 6) highlights the total number of users who succeeded (within or outside of the benchmark time) in each test task. Users performed best in Tasks 1 and 4 with a success rate > 70%. Task 2 was performed quite well with a success rate of 60%, however in contrast, users completed Task 3 poorly with only a 23.33% success rate. Tasks 5 and 6 ranked at the bottom, with failure rates just above 83%, demonstrating that they were the most difficult tasks. Two possible reasons for the poor performance of tasks 3, 5 and 6 are that 1) the tasks were too difficult for the users to achieve or 2), the application was experiencing a very low level of

usability. The success rate, time, number and type of problems will help to clarify the real reason, which will be discussed in the following sections. However, it is worth mentioning that the researchers carefully followed a proven methodology taken from the literature review, and ensured that all the test tasks were set up to emulate realistic situations as far as possible. Also the researchers, prior to each task, made sure that each user understood the task and that he/she fully was happy to perform it, with no doubts at all. This was also supported by the pilot study which enabled the researchers to fully address the participants' comments and doubts. Finally, no user pointed out any difficulty in the task, or the criteria of each task, during or after each task. The average success rate for the tested application can be calculated using the following formula:

$$\frac{\text{Number of tasks completed successfully}}{\text{the total number of task performed}} *100$$

By applying the above formula we obtain the following results:

$$\text{Success rate} = 80/180*100= 44.44\%$$

5.1.2. Time on Tasks

	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Total	Average
Total Time	90.54	74.37	99.27	37.15	184.44	115.25	601.02	100.17
Maximum Time	07:05	06:54	06:31	03:01	10:55	06:17		
Minimum Time	00:59	00:50	01:20	00:28	01:51	01:10		

TABLE 7: Tasks Total, Maximum and Minimum Times.
Times are in minutes and seconds (mm:ss).

Table 7 above shows the total time spent by all the participants on each task. For example, the total time spent by all the users on Task 1 was 90 minutes and 54 seconds. Again, it is noticeable that Tasks 5 and 6 took the longest total time to complete, a fact which correlates with the high number of problems encountered by all users in these two tasks. The actual total number of problems in Task 5 was 234, and in Task 6 was 132; this reflects the significant usability problems uncovered by these two tasks (see section 5.4 and Figure 8).

5.2. Participant's Satisfaction

5.2.1. System Usability Scale (SUS)

Satisfaction shows how pleased people are with their experience using the tested application (also referred to as user experience). This was measured with a 10 question survey (APPENDIX B), taken at the end of the test using a marking system called the System Usability Scale (SUS). Each question has a scale from 1 (Strongly disagree) to 5 (strongly agree). From these surveys a SUS Score is derived, which indicates how user-friendly participants perceive the application to be. A SUS Score can range from 0 (worst) to 100 (best) [42]. Bangor, Kortum and Miller [43] have developed a seven-point scale, adding a user-friendliness item to the 10 questions of SUS (see Figures 4 & 5). They produced a grading scale in which SUS scores below 60 are graded as 'F', between 60 and 69 as a 'D', between 70 and 79 as a 'C', between 80 and 89 as a 'B', and 90 and above are granted an 'A' (see Figure 3). The overall SUS Score of the tested application Souq.com was 50.42, which achieved a usability grade of 'F', signifying that the usability level is far below that expected.

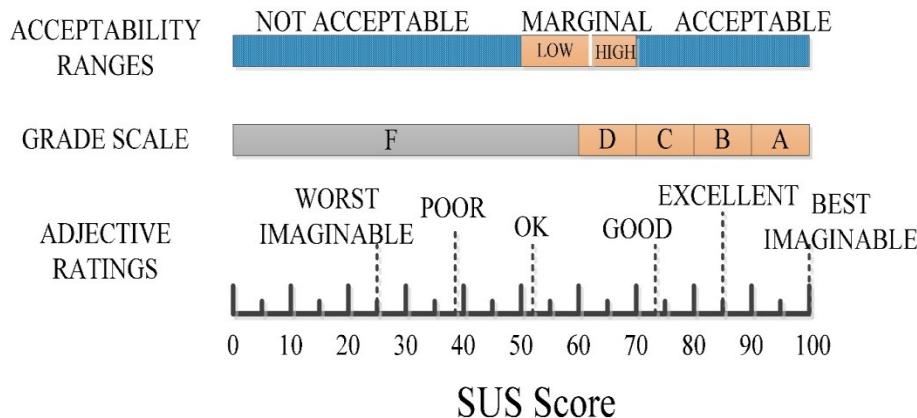


FIGURE 3: A comparison of the adjective ratings, acceptability ranges and grading scales in relation to SUS score. Adapted from [43].

Figure 4 below presents the average rating for each SUS question. Due to the wide use of SUS in usability studies, Sauro [44] conducted an extensive study, reviewing the existing research on SUS and analysing the data from more than 5,000 users across approximately 500 different evaluations. He concluded from the data that SUS is a reliable and valid measure of perceived usability. Furthermore it performs as well as, or better than, commercial questionnaires and home-grown internal questionnaires. In statistics, Cronbach's α alpha (alpha) measures the coefficient of internal consistency. It is normally used as an estimation of the reliability of a psychometric test for a sample of examinees [45]. The value of Cronbach's alpha should be above 0.70, although a value of 0.8 or higher is preferred and a slightly lower score of 0.60 is acceptable [46]. Therefore, the researchers conducted Cronbach's Alpha reliability test using IBM SPSS for the 10 items of the SUS questionnaire, and obtained the following result (Table 8):

Reliability Statistics	
Cronbach's Alpha	N of Items
.856	10

TABLE 8: Cronbach's Alpha.

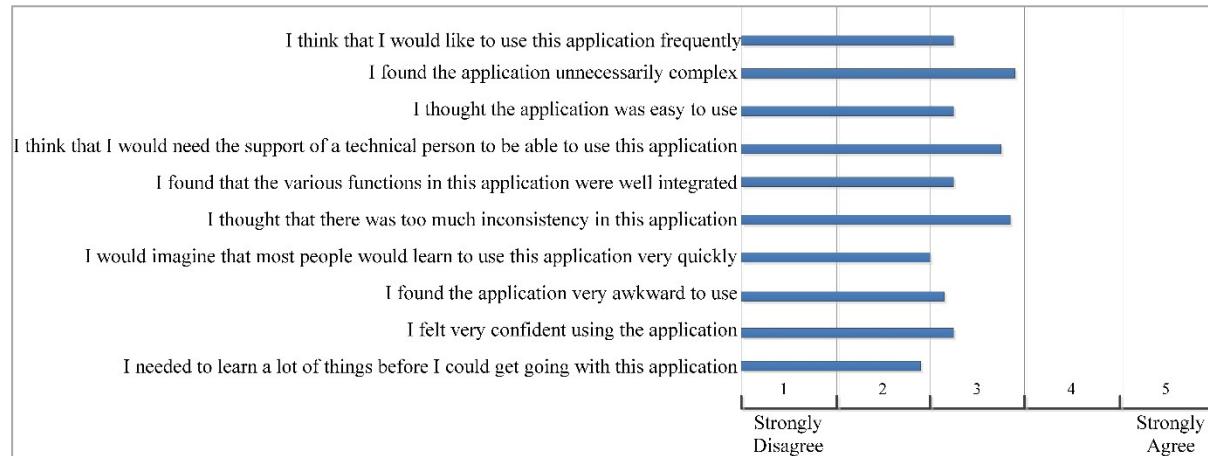


FIGURE 4: Users' answers to SUS questionnaire.

From the graph above, it can be summarized that the users agree with the negative statements more than with the positive statements. For example, none of the users agreed with the statement "I think that I would like to use this application frequently", indicating that they are

unlikely to use this mobile application (m-transaction) again. This leads to the conclusion that the poor usability of m-transactions negatively affects the users' intention to use them in the future. Thus the low usability of such applications as the one tested can answer the question "Why does the ease of use of mobile transactions have the highest impact on the intention to use mobile transactions in Saudi Arabia?".

Furthermore, a correlation between the total numbers of problems for each user with the total time for each person, using Pearson Correlation test, gave the following result (Table 9):

Total Problems & Total Time	Total Problems & SUS
0.383	-0.001

TABLE 9: Pearson Correlation.

The above table shows that there is a statistically strong positive relationship between the total time spent and the number of problems discovered during the tasks. It implies that the more problems the user discovers, the more time the user spends on each task. This should be expected as a logical result. The correlation test reveals that the relationship between the total number of problems encountered and the users' satisfaction is negative, which means the more problems discovered by the user, the less satisfied they felt with the system. Interestingly, this result is in contrast with the findings of several studies, e.g. [47], which highlights that subjective evaluation techniques such as questionnaires should be interpreted with caution when used with participants from the Eastern culture, as they tend to rate the system under evaluation positively despite clear indication of a poor performance, in order to show politeness to the evaluator. However, in this study it was found that Saudi participants were not influenced by such a culture effect.

5.2.2. Adjective Rating Scale

Although a SUS questionnaire is considered an effective tool for assessing the usability of products including web sites, mobile phones and any product or system that has an interactive face, there is still a need for information describing how the numeric score translates into an absolute judgment of usability. Bangor, Kortum [43] have added an adjective rating scale to the SUS Score in order to help interpreting the numeric value of the SUS Score and explain the results to non-human factors professionals (see Figure 5). This study has applied this addition.

Overall, I would rate the user-friendliness of this application.

Worst Imaginable	Awful	Poor	OK	Good	Excellent	Best Imaginable
<input type="radio"/>						

FIGURE 5: The adjective rating statement appended to the SUS Score.

Figure 6 below illustrates the participants' overall judgment of the tested application. It can be seen that the majority of users were not happy with the application's usability level. More than half of the users (53.32%) had a negative user experience, with 36.66% of the users rating it as 'Poor', 13.33% describing it as 'Awful', and 3.33% assessing it as 'Worst Imaginable'. On the other hand, 36.66% of the participants commented on their experience as 'OK'. 10% chose 'Good' to reflect their opinion of the tested application, however none of the participants rated it as 'Excellent' or 'Best Imaginable' (see Figure 6).

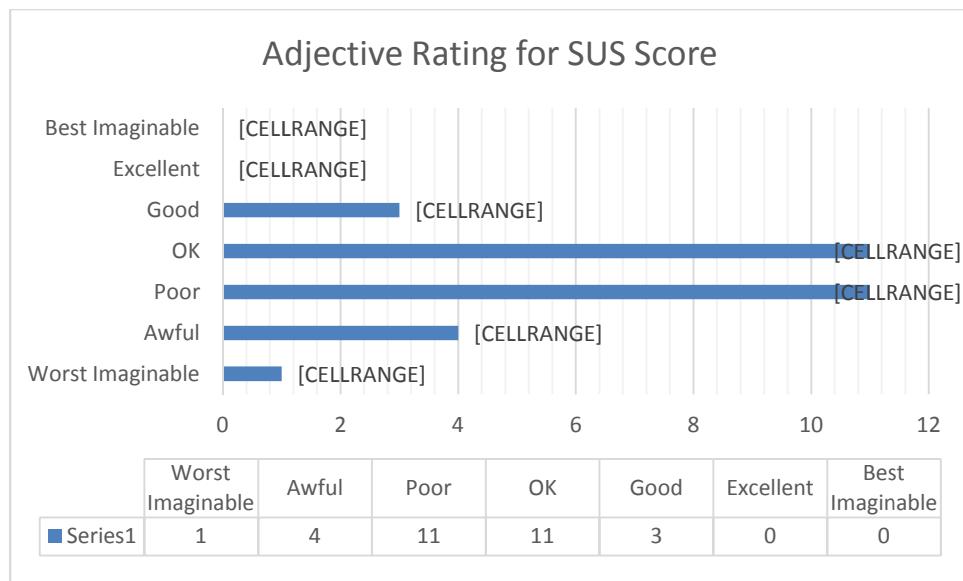


FIGURE 6: Overall users' judgment on the tested application.

5.3. Usability Problems Discovered

5.3.1. Number of Problems Discovered

In Task 1, which is about registering with the application, users discovered 92 problems in total. During Task 2, when participants had to find a smart phone and add it to the basket, they encountered 11.66% of the total problems. They identified 14.61% of the total problems in Task 3 while trying to find a book that they liked and add it to the basket. Users faced 67 usability problems when finding a football as a gift in Task 4, and in Task 5, where they had to conduct a transaction for one of the products inside the basket, they found most of the usability problems, a total of 234. This meant that Task 5 alone accounted for more than 32% of the total usability problems experienced by the users. Finally, the users had to contact the support team to ask about the arrival/delivery time of their purchased product in Task 6, which generated 132 (18.54%) usability problems. In total, 712 individual usability problems were observed (see Figure 7).

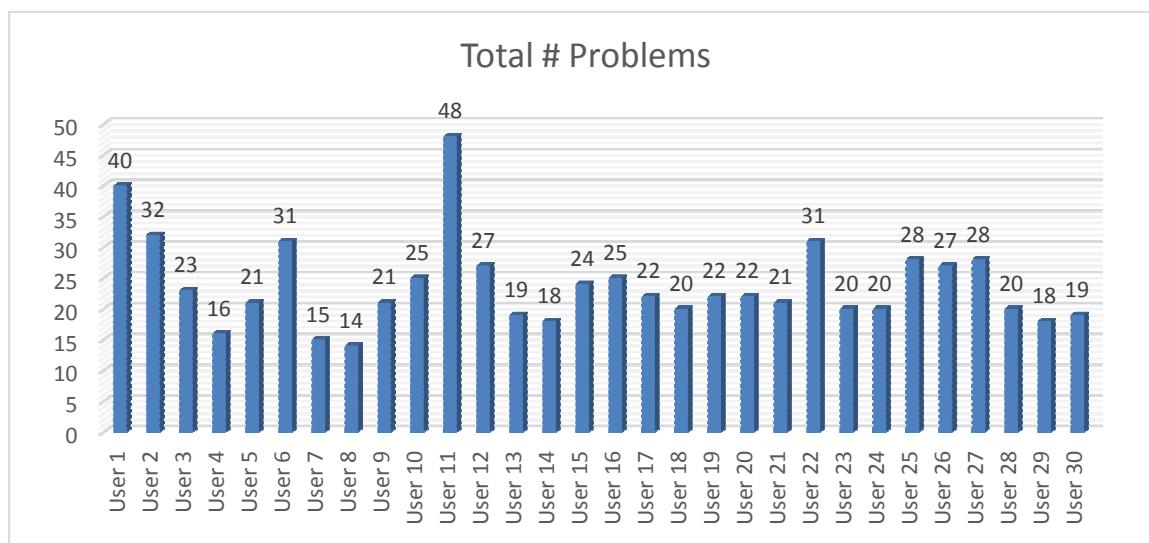


FIGURE 7: Users' total number of problems.

The above graph shows the total number of usability problems discovered by each user over the 6 tasks. The average number of problems amongst users is 24, and the median is 22. This again

supports the claim that conducting transactions via mobile devices is still in its infancy stage, and that a lot of effort is needed to improve the users' experiences.

As mentioned above, the total number of usability problems discovered by the 30 users across the 6 tasks was 712. However, the researchers used Excel software (a product of the package Microsoft Office Professional Plus 2013) to help with filtering and organizing the final problems as follows:

- 1- The main author transcribed all the problems discovered by the users from the observation sheet into one Excel sheet.
- 2- Repeated problems were consolidated into a single problem.
- 3- Two independent experts categorised all the detected problems into 8 specific problem categories.
- 4- The final number of problems in each category was calculated.

This process reduced the total number of problems from 712 to 136, thus strengthening the classification of the types of problems discovered. This is an approach similar to that used by [29, 48] in which two usability experts were asked to collaboratively categorize all detected problems into eight specific problem types. Hornbaek and Frøkjeer [49] suggest that team classification/matching is a promising technique to strength the reliability of usability problems data. The two experts were also asked to rate collaboratively the severity of the uncovered problems. With an average score of 2.55 for all problems, this indicates that the majority of the test applications had major problems.

These categories are as follows:

- **Comprehension:** The participant finds that the information in the application is not clear or does not make sense, which usually confuses the user when completing the task or makes it difficult to make the right decision.
- **Relevance:** in this category the user finds unnecessary information and details that may make the task more difficult to accomplish.
- **Completeness:** This category involves problems such as the lack of required information, guidance or help, lack of functions, options or features.
- **Design/Layout:** This category refers to where the user may face problems or difficulties with regard to the design, the page structure and layout, styling and formatting, language used, and navigation.
- **Correctness:** This category groups together problems such as a violation of syntax, unresponsive functions, links and buttons, or the application failing to do what it is supposed to do and crashing.
- **Data entry:** This category encompasses the problems experienced by the users when entering data into the application. This includes the use of GUI (Graphical User Interface) widgets, the virtual smart phone keyboard, editing, clicking and filling in forms and text/number fields.
- **Visibility:** This category includes problems such as not being able to find a particular link, button, piece of information or a particular page.
- **Download delays:** This category summarizes the problems regarding the loading time for pages and processes to be completed.

5.3.2. Types of Usability Problems Discovered

Tables 10 and 11 illustrate the proportion and number of problems identified in each category and the number of total problems. The actual final problem descriptions and their categories can be seen in APPENDIX D.

Category	# Final Problems
Relevance	4
Data entry	7
Design/Layout	46
Completeness	25
Visibility	9
Download delays	8
Correctness	23
Comprehension	14
Total	136

TABLE 10: Problem Categories.

According to [50], usability problems can fall into one of the following categories of severity: not a usability problem, cosmetic, minor, major, and catastrophic. To ensure an objective assessment of the problems discovered in this study, the researchers sent the final set of problems identified by the participants to two usability experts with many years of experience in the usability field, who then classified their severity based on their frequency, impact and persistence. Table 11 below shows the frequency of each severity category.

Problem Types	No.
Cosmetic Problems	13
Minor Problems	51
Major Problems	56
Catastrophic Problems	16
Total Problems	136

TABLE 11: Problems Types.

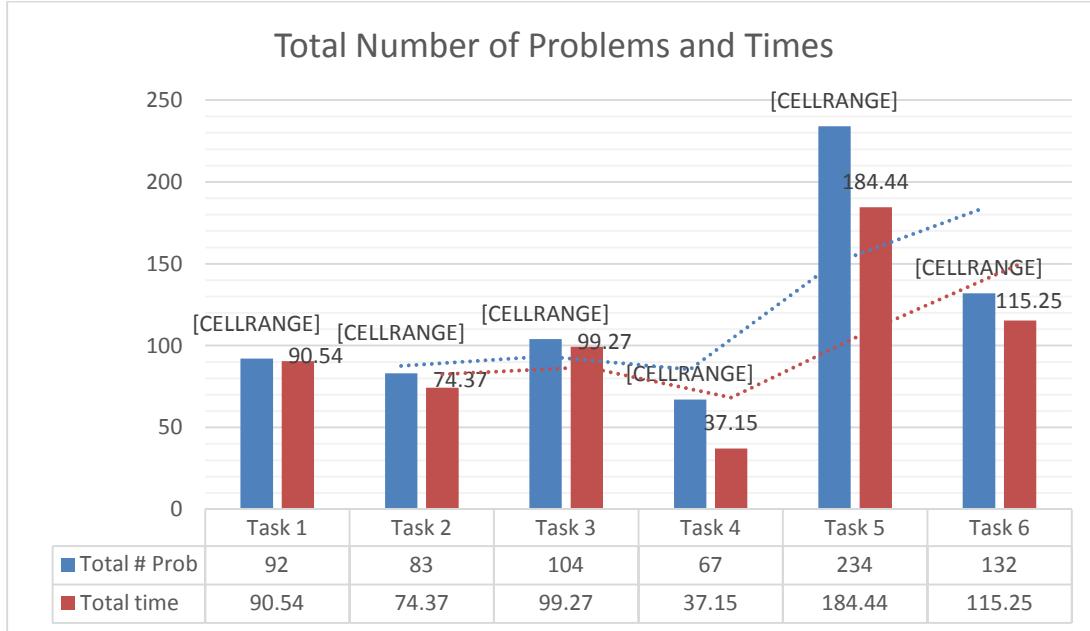


FIGURE 8: Total number of problems in conjunction with total time.

The above chart (Figure 8) summarises the total number of final problems discovered in each test task by the users. It also shows the total time spent on all 6 tasks. As expected, it highlights the fact that when the users encounter more problems, they spend more time completing the task. Tasks 5 and 6 took the longest total time, and users encountered the highest number of problems when completing them; the problems clearly cost them time. This was mirrored in Task 6 which had the second highest number of problems and the second longest completion time. This, furthermore, leads to the conclusion that users encounter more usability problems when using a mobile device (compared to PC) and, therefore spend more time completing the task. Using the adjusted Wald formulae (also called the modified Wald interval), we can estimate the confidence interval for the completion of the fifth task by entering the number of those who successfully passed this task and the total number of participants [51]. The results show that we can be 95% confident that the actual completion rate for this task is between 6% and 34% of the whole population. In other words, we can be 95% confident that only between 6% and 34% of the whole targeted population (i.e. Saudi users) would be able to complete this task. This illustrates the poor usability of the app and how difficult it can be for Saudi users to conduct m-transactions.

6. A COMPARATIVE EVALUATION

The following provides a comparative evaluation of the research finding with other research findings to reflect other similar studies carried by other authors in the same field to justify the novelty and authenticity of the research work. Öztürk and Rızvanoğlu [22] conducted a mobile usability tests on three different private shopping applications/mobile sites in Turkey to understand the importance of usability with regards mobile commerce. They applied a multi-method approach in testing the underlying application (i.e. think aloud, eye-tracking and video recording). The sample size for the experiment was 11 Turkish senior year university students, who have had mobile internet experience and consider a potential user of private shopping mobile applications. The findings of the experiment revealed that there are significant mobile usability problems with regards to homepage logging in, navigational browsing, results listing, product detail and forms and checkout. They concluded that: "despite the growing interest towards m-commerce platforms, poor usability of mobile sites and applications for commerce activities stands out as a major obstacle for the slow adoption of mobile solutions." [22]. Therefore, their findings in consistent with this research results. This also applies to the study by [52] who conducted an exploratory study of usability issues for ten consumer wireless Web sites across three form factors in USA: Palm OS based wireless PDAs (Personal Digital Assistant), Internet-enabled WAP (Wireless Application Protocol) phones, and Windows CE based Pocket PCs. They used cognitive walkthrough and heuristic evaluation methods to evaluate the usability of wireless sites. Their findings revealed a number of usability problems (e.g. long download, broken connection, vertical and horizontal scrolling, information overload, depth of site structure, and search).

7. CONCLUSION AND RECOMMENDATION

This research sheds some light on usability levels with respect to m-transactions in Saudi Arabia. The outcomes of this study may contribute to the market stakeholders' understanding of their potential customers' needs and concerns. Exploring the market, especially at the time when m-transactions are still in their infancy in most countries around the globe, is critical for industry stakeholders if they are to ensure the success of this emerging market. Therefore, this research highlights the most important factor (i.e. usability) that needs to be considered in order to support the proliferation and advancement of m-transactions in developing countries, and especially in Saudi Arabia.

Although the results show that the usability level of the tested mobile application is remained unsatisfactory, there are a number of justifiable reasons. One reason is due the lack of experience of applying such a new technology within the context of Saudi Arabia from both parties (the application stakeholder and developer and the Saudi users). This because m-transaction is still in its infancy stage and yet to reach the maturity level. Furthermore, the application developer and the stakeholder should consider paying attention in increasing the

Saudi user's awareness of the underlying technology. Provide clear and enjoyable tutorials (e.g. video tutorial or interactive web based tutorial). Consider the problems that discussed in the recommendations. Consider conducting usability testing of the application with a reasonable number of Saudi users before releasing/launching it into the Saudi market. The government, moreover, can help in increasing the awareness in TV shows and news. Provide facilities and promotions to companies interested in adopting m-transaction. Integrate m-transaction on the online government services that already exist (e.g. the electronic services of the Ministry of Interior, the Passport General Department and the Ministry of Labor ...etc).

With reference to the results of the experiments, this study was able to generate a visualisation of the current situation of m-transaction usability which shed light on:

- The current usability level of m-transactions within the Saudi context.
- The number and types of problems that Saudi users may encounter while conducting m-transactions, thus deepening understanding and awareness of these issues.
- Saudi users' performance and rate of success with regards to conducting m-transactions.

All of the above provide opportunities and the potential for policy makers, businesses and application developers to develop and personalize the m-transactions process, so that it fits Saudi users' requirements. This ultimately can have a significant indirect impact on the intention to use m-transactions in Saudi Arabia and therefore, the adoption and use of m-transactions from a consumer's perspective. Mobile application developers and designers, companies extending their business into mobile commerce, and governments, need to pay real attention to the m-transaction usability levels shown in the tables and figures in this study. Failure to do so could lead to an increase in the number of Saudi users neglecting or avoiding the use of m-transactions. The results of this study could be further validated using other usability methods (e.g. interviews) and could also be applied to a larger number of Saudi users or to different locations of the kingdom to confirm the current results. Also, including female participants in a future study could help to generalize the results.

Although the tested application Souq.com enjoys great popularity and represents realistic m-transaction functionalities, the results clearly showed that it was unsatisfactory as a usable application. It has a very low SUS Score (50) compared to 60 which is recommended as an acceptable score. It also ranked in Bangor et al.'s classification as 'F' which is considered as unacceptable. Furthermore, it has a very low success rate and efficiency. Therefore, the designers and developers of this application have to pay more attention to the usability level of their application and consider the usability problems discovered through the experiments, especially to those categories with a higher number of problems such as 'design/layout', 'completeness', 'correctness' and 'comprehension'. In more detail, the problems and issues that need to be considered are: (1) simplifying the forms (e.g. registration, contact us and add an address); (2) clearly labelling the buttons with names that are easy to understand and reflect the actual function of that button; (3) redesigning the filter function in all sections in a simple way (e.g. so there is no need to press 'apply' for every specification for finding a product); (4) managing the basket needs to be improved (e.g. the product amount does not handle the value of 0 (zero) and the button 'add to the basket' needs to be renamed); (5) adding further functions (e.g. the ability to delete more than one item at a time); (6) correcting errors of logic (e.g. the authors' names were written in two different languages, English and Arabic, the challenge question was written in Arabic yet does not accept the answer while the keyboard is Arabic, and the whole payment confirmation page was written in English while the chosen interface language was Arabic); and (7) removing the distracting popup messages (e.g. after clicking contact us or while looking for a product). A usability test should be conducted prior the launching of any mobile application to ensure that the usability level is satisfactory. Future research has to pay attention to the importance of usability to consumers' acceptance of IT innovations such as m-transactions.

With regards to limitations, in a conservative society such as Saudi, the authors were not able to involve females, therefore, this study has considered only male participants and another study

could involve the opposite gender to get more generalizable/interesting results. When another suitable mobile application becomes available, a further study could test that application and compare the results with our study.

8. NOVELTY

The extensive literature review conducted by the researchers revealed that there have been no studies which test a mobile application within the context of Saudi Arabia, neither do any shed light on the importance of usability as an important factor for accepting or using any new IT innovation from Saudi users' perspective. Therefore, this research is considered as a first of its kind in Saudi Arabia.

9. FUTURE RESEARCH DIRECTIONS

The future research directions would consider to test any new m-transaction mobile application in the Saudi market and compare the results with the current study's findings. Also a usability test for a newer version of Souq.com would be interesting.

10. REFERENCES

- [1] J. H. Cheong & M.-C. Park, 'Mobile Payment Adoption in Korea: Switching from Credit Card'. In *15th International Telecommunication SocietyEurope Regional Conference*, 1-22, 2008.
- [2] T. Dahlberg, N. Mallat, J. Ondrus & A. Zmijewska, 'Past, present and future of mobile payments research: A literature review'. *Electronic Commerce Research and Applications*, 7, 165-81, 2008.
- [3] O. Bamasak, 'Exploring consumers acceptance of mobile payments – an empirical study'. *Int J Information Technology, Communications and Convergence*, 1, 173-85, 2011.
- [4] M. Alqahtani, A. Al-Badi & P. Mayhew, 'Exploratory Study of M-Transaction: User's Perspectives'. *The Electronic Journal of Information Systems in Developing Countries*, 60, 1-22, 2014.
- [5] M. Alqahtani, A. Al-Badi & P. Mayhew, 'The Enablers and Disablers of E-Commerce: Consumers' Perspectives'. *The Electronic Journal of Information Systems in Developing Countries EJISDC*, 54, 1-24, 2012.
- [6] M. Alqahtani, R. S. Alroobaea & P. J. Mayhew, "Building a Conceptual Framework for Mobile Transaction in Saudi Arabia: A User's Perspective". In: Science and Information Conference (SAI), 2014 London. IEEE, 967-73, 2014.
- [7] D. Matthew, 'Factors that impact consumer adoption of mobile banking within the m-commerce domain'. 2014.
- [8] Q. Min & S. Li, 'From Usability to Adoption - A New M-commerce Adoption Study Framework'. *Communications and Mobile Computing, 2009 CMC '09 WRI International Conference on*, 3, 309-13, 2009.
- [9] J. Buranatrived & P. Vickers, "An Investigation of the Impact of Mobile Phone and PDA Interfaces on the Usability of Mobile-Commerce Applications". In: Proceedings of IEEE International Workshop on Networked Appliances, 2002 Liverpool. IEEE, 90-5, Year.
- [10] V. Venkatesh, V. Ramesh & A. P. Massey, 'Understanding usability in mobile commerce'. *Commun ACM*, 46, 53-6, 2003.
- [11] Y.-M. Li & Y.-S. Yeh, 'Increasing trust in mobile commerce through design aesthetics'. *Computers in Human Behavior*, 26, 673-84, 2010.

- [12] D. Zhang & B. Adipat, 'Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications'. *International Journal of Human-Computer Interaction*, 18, 293-308, 2005.
- [13] J. Nielsen & R. Budiu. *Mobile usability*, Pearson Education, 2012.
- [14] A. Holzinger, 'Usability engineering methods for software developers'. *Commun ACM*, 48, 71-4, 2005.
- [15] J. Kjeldskov & J. Stage, 'New techniques for usability evaluation of mobile systems'. *International journal of human-computer studies*, 60, 599-620, 2004.
- [16] F. Nayebi, J. M. Desharnais & A. Abran, "An Expert-Based Framework for Evaluating iOS Application Usability". In: Software Measurement and the 2013 Eighth International Conference on Software Process and Product Measurement (IWSM-MENSURA), 2013 Joint Conference of the 23rd International Workshop on, 23-26 Oct. 2013 2013. 147-55, Year.
- [17] T. Kallio & A. Kaikkonen, 'Usability testing of mobile applications: A comparison between laboratory and field testing'. *Journal of Usability studies*, 1, 23-8, 2005.
- [18] R. Alroobaea, A. Al-Badi & P. Mayhew, 'Generating a Domain Specific Inspection Evaluation Method through an Adaptive Framework: A Comparative Study on Educational Websites'. *International Journal of Human Computer Interaction (IJHCI)*, 4, 88, 2013.
- [19] R. Alroobaea, A. H. Al-Badi & P. J. Mayhew, 'Generating a Domain Specific Inspection Evaluation Method through an Adaptive Framework'. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 4, 2013.
- [20] F. Nayebi, J.-M. Desharnais & A. Abran, "The state of the art of mobile application usability evaluation". In, 2012 2012. 1-4, Year.
- [21] J. Nielsen. 1995. *Severity Ratings for Usability Problems* [Online]. City. Available: <http://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/> [Accessed 25-09-2014 2014].
- [22] Ö. Öztürk & K. Rızvanoğlu, M-Commerce usability: an explorative study on turkish private shopping apps and mobile sites. *Design, User Experience, and Usability Web, Mobile, and Product Design*. City: Springer, 2013.
- [23] S. Yang, Y. Lu, S. Gupta, Y. Cao & R. Zhang, 'Mobile payment services adoption across time: An empirical study of the effects of behavioral beliefs, social influences, and personal traits'. *Computers in Human Behavior*, 28, 129-42, 2012.
- [24] C. Kim, M. Mirusmonov & I. Lee, 'An empirical examination of factors influencing the intention to use mobile payment'. *Computers in Human Behavior*, 26, 310-22, 2010.
- [25] P. G. Schierz, O. Schilke & B. W. Wirtz, 'Understanding consumer acceptance of mobile payment services: An empirical analysis'. *Electronic Commerce Research and Applications*, 9, 209-16, 2010.
- [26] K. Petrova & R. Mehra, "Mobile Payment: An Exploratory Study of Customer Attitudes". In: Wireless and Mobile Communications (ICWMC), 2010 6th International Conference on, 20-25 Sept. 2010 2010. 378-83, Year.
- [27] G. Wei, Z. Xinyan & M. Yue, 'Literature review on consumer adoption behavior of mobile commerce services'. *E -Business and E -Government (ICEE), 2011 International Conference on*, 1-5, 2011.

- [28] O. Alhadreti, R. Alroobaea, K. Wnuk & P. J. Mayhew, "The Impact of Usability of Online Library Catalogues on the User Performance". In, 2014 2014. IEEE, 1-4, Year.
- [29] M. J. Van Den Haak, M. D. T. De Jong & P. J. Schellens, 'Evaluating municipal websites: A methodological comparison of three think-aloud variants'. *Government Information Quarterly*, 26, 193-202, 2009.
- [30] R. Benbunan-Fich, 'Using protocol analysis to evaluate the usability of a commercial web site'. *Information & Management*, 39, 151-63, 2001.
- [31] Souq.Com. 2014. *About us* [Online]. City: Saudi.souq.com. Available: <http://saudi.souq.com/sa-en/about-us/c/> [Accessed 03/09/2014 2014].
- [32] A. A. Alrobai, R. Al-Badi & P. Mayhew, 'Investigating the usability of e-catalogue systems: modified heuristics vs. user testing'. *Journal of Technology Research*, 2012.
- [33] C. Snyder. *Paper prototyping: The fast and easy way to design and refine user interfaces*, Morgan Kaufmann, 2003.
- [34] M. Maguire, 'Context of use within usability activities'. *International Journal of Human-Computer Studies*, 55, 453-83, 2001.
- [35] C. M. Barnum. *Usability testing essentials: ready, set... test!*, Elsevier, 2010.
- [36] J. S. Dumas & J. Redish. *A practical guide to usability testing*, Intellect Books, 1999.
- [37] J. Sauro & J. R. Lewis. *Quantifying the user experience: Practical statistics for user research*, Elsevier, 2012.
- [38] J. Sauro & E. Kindlund, "How long should a task take? identifying specification limits for task times in usability tests". In, 2005 2012. Year.
- [39] J. R. Lewis, 'Sample sizes for usability tests: mostly math, not magic'. *interactions*, 13, 29-33, 2006.
- [40] R. Alroobaea & P. J. Mayhew, "How Many Participants Are Really Enough for Usability Studies?". In: Science and Information Conference (SAI), 2014 London. IEEE, 757-67, 2014.
- [41] A. Bangor, P. T. Kortum & J. T. Miller, 'An empirical evaluation of the system usability scale'. *Intl Journal of Human–Computer Interaction*, 24, 574-94, 2008.
- [42] J. Brooke, 'SUS-A quick and dirty usability scale'. *Usability evaluation in industry*, 189, 194, 1996.
- [43] A. Bangor, P. Kortum & J. Miller, 'Determining what individual SUS scores mean: Adding an adjective rating scale'. *Journal of usability studies*, 4, 114-23, 2009.
- [44] J. Sauro. 2011. *Measuring Usability With The System Usability Scale (SUS)* [Online]. City: Measuring Usability LLC Main Offices. Available: <http://www.measuringusability.com/sus.php> [Accessed 24-09 2014].
- [45] L. J. Cronbach, 'Coefficient alpha and the internal structure of tests'. *psychometrika*, 16, 297-334, 1951.
- [46] J. Hair, W. Black, B. Babin & R. Anderson, 'Multivariate data analysis: a global perspective: Pearson Education'. 2010.

- [47] L. Herman, "Towards effective usability evaluation in Asia: Cross-cultural differences". In, 1996 1996. IEEE, 135-6, Year.
- [48] M. Alshamari & P. Mayhew, "Task design: its impact on usability testing". In: The Third International Conference on Internet and Web Applications and Services, 08-13 June 2008 Athens. IEEE, 583–9, 2008.
- [49] K. Hornbaek & E. Frøkjeer, 'A Study of the evaluator effect in usability testing'. *Human-Computer Interaction*, 23, 251-77, 2008.
- [50] J. Nielsen, 'Usability Engineering'. 1st Ed Cambridge: Morgan Kaufmann, 1993.
- [51] J. Sauro. 2005. *Confidence Interval Calculator For A Completion Rate* [Online]. City: Measuring Usability LLC. Available: <http://www.measuringu.com/wald.htm> [Accessed 09-10-2014 2014].
- [52] S. S. Chan, X. Fang, J. R. Brzezinski, Y. Zhou, S. Xu & J. Lam, 'Usability for Mobile Commerce Across Multiple Form Factors'. *J Electron Commerce Res*, 3, 187-99, 2002.

APPENDIX A

The tasks used in this study:

Task 1: < Registering with the Application > You are a new user of the Souq.com application and want to be a registered user, so that you can use all the application functions. What do you do?	
Goals / output:	Register successfully into the app, resulting in full use of all the app's functions. To find out to what extent the registering interfaces, processes and functions are easy to use.
Inputs / Assumptions:	Names (first and last) Gender Country Email address User Name Password Confirmation of Password Terms Agreement Confirmation of Email The user should have a valid email address Task 1 is completed
Steps:	Go to the Homepage Complete the registration process (several fields/steps) Log in View the application interface and its services (functions).
Time for expert:	01:19 minutes.
Instructions for user:	Please open the app in your smart phone and register as a new user, and then try to navigate through its services and functions.
Notes:	The user can sign up a new email if he/she does not have one already. The users' overall feedback about this task will be recorded/noted. When you find the product that you want to buy, please add it to the cart and go to the homepage so you can continue shopping.
Task 2: < Find a Product to Purchase “Smart Phone” > You have just decided to get a new mobile phone and you want to search for and buy a new smart phone. What do you do?	
Goals / output:	Successfully locating a product to purchase, resulting in adding the product to the cart.

	To find out to what extent finding (Searching, navigating, filtering and browsing) the required information/product is easy.
Inputs / Assumptions:	Product specifications details (several steps) Has confirmation email as a result of task 1 Each user chooses one product to purchase (Smart phone) Completed tasks 1
Steps:	Go to the Homepage Go to Mobile and Accessories category Choose Mobiles Put in the filter the product specifications. Scroll down until locating the product Add it to the cart
Time for expert:	01:30 minutes
Instructions for user:	Please find a smart phone to purchase that has these specifications: Samsung Galaxy S4, New, Blue, 16GB, 1,799SAR.
Notes:	The user may not find an option to choose the colour, condition, data storage and price so he/she has to overlook these steps. The users' overall feedback about this task will be recorded/noted. When you find the product that you want to buy, please add it to the cart and go to the homepage so you can continue shopping.
Task 3: < Find a Product to Purchase "Book" >	
Goals / output:	Successfully identifying a product to purchase, resulting in adding the product to the cart. To find out to what extent finding (different categories, functions, filters and interfaces and navigational structure) the required information/product is easy.
Inputs / Assumptions:	The book specifications. Tasks 1 and 2 were accomplished
Steps:	Go to the Homepage Go to browse all categories Choose books section Put in the filter the product specifications. Scroll down until locating the product Add it to the cart
Time for expert:	00:36 minutes
Instructions for user:	Please find a book to purchase. Please make sure that the book has the following specifications: Name: "سبنسر" "الاعتماد على الذات" type: "من حرك قطعة الجبن الخاصة بي" Author: "سبنسر" "الاعتماد على الذات" publisher: "مكتبة رير" "العربية" Language: "العربية", and then add it to your cart.
Notes:	The user needs to consider the filters in this task. When you find the product that you want to buy, please add it to the cart and go to the homepage so you can continue shopping. The users' overall feedback about this task will be recorded/noted.
Task 4: < Find a Gift to Purchase "Football" >	
Your young brother has graduated recently. You want to search for and buy a football for him as a gift. What do you do?	
Goals / output:	A successful finding of a low-price product, resulting in adding the product to the cart. To find out to what extent finding the required information/product is easy (e.g. controlling the results shown and using search functions).
Inputs / Assumptions:	Logged in The product is available and in stock Previous tasks were successfully completed
Steps:	Go to the homepage Locate the cursor inside the search bar Type in "كرة بيد سبيس" Scroll down to locate the product Choose a product with 8SAR of cost Add it to the cart
Time for	00:46 minutes

expert:	
Instructions for user:	Please use the search bar to find a football called "کری بیرد سبیس" and make sure the price is 8SAR. Add this product to your cart.
Notes:	There are a number of similar products; the user has to choose the one that costs 8SAR. The users' overall feedback about this task will be recorded/noted.
Task 5: < Managing the Cart and a Payment Transaction > You need to remove all the items from your cart except the item that cost 8 SAR and then conduct the payment. What do you do?	
Goals / output:	A successful control of the cart, resulting in conducting the payment transaction. To find out to what extent managing the cart and conducting the payment transaction is easy.
Inputs / Assumptions:	A number of items are already in the cart Credit card
Steps:	Go to the Homepage Click the right top button Choose my cart Remove all the items except the item that cost 8SAR. Conduct the payment (several steps)
Time for expert:	02:30 minutes
Instructions for user:	Please go to your cart and check the products added to your basket. Remove all the products except the product that you added last (which cost 8SAR). Purchase this product by conducting the payment transaction.
Notes:	The users' overall feedback about this task will be recorded/noted.
Task 6: < Managing the Cart and a Payment Transaction > You have a request about the delivery time for your purchased product. What do you do?	
Goals / output:	A successful use of "Contact us" function, resulting in sending the request. To find out to what extent sending a request about a product is easy.
Inputs / Assumptions:	Has the Order ID. The product was already purchased in task 5.
Steps:	Go to "more" Click on "Contact us" Put your email address Chose the request type Type in the request in the text box Answer the challenge question Click the send button
Time for expert:	02:20 minutes
Instructions for user:	Assume you have a request about the delivery time for the product you just purchased. Please send a request to the support team via "Contact us".
Notes:	The user will send the request about an already purchased product if he/she did not succeed in task 5.

APPENDIX B

Introductory Page

User Questionnaire

Information

Thank you very much for taking the time to participate in this research. This study is a part of a PhD research and the participation of this usability test/experiment will significantly contribute to the success of this research. This study aims to evaluate the current usability level of a mobile website/application “Souq.com – سوق.كوم” from the user perspective in Saudi Arabia. Please give your opinion and/or preferences honestly and frankly; there are no right or wrong answers.

Outline

This interview includes two main sections about the underlined mobile website/application. A first section is about the user background and, the second section is a multiple choice questions to illustrate your satisfaction level about the underlined website/application.

Confidentiality

Please be assured that any information provided will not be used for any commercial purpose and no one will have access to this information apart from the researchers. Also please feel totally safe, that any information provided will be treated anonymously in the strictest confidence and none of the participants will be individually identifiable in the resulting thesis, reports or other publications.

These data will be retained securely in the University of East Anglia premises for ten years and after that it will be destroyed. You are, of course, entirely free to discontinue your participation at any time or to decline to answer a particular question.

Thank you again for taking the time to respond to this interview. If you have any further questions please contact the researcher (Mr. Mohammed Alqahtani) by email at m.alqahtani@uea.ac.uk or the project supervisor Dr Pam Mayhew at p.mayhew@uea.ac.uk.

*Mr. Mohammed Alqahtani
School of Computer Science,
University of East Anglia*

Norwich, NR4 7TJ

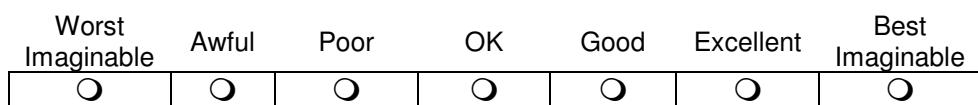
Section A: User Profile

1. Overall, how do you rate your mobile skills?
 - No Skills
 - Beginner
 - Intermediate
 - Advanced
2. How long have you been using mobile internet?
 - Less than a year
 - 1-2 years
 - 3-5 years
 - More than 5 years
3. Have you visited "Souq.com – سوق.كوم"?
 - Yes
 - No
4. If yes, how often do you use "Souq.com – سوق.كوم"?
 - Every day
 - Weekly
 - Monthly
 - Hardly ever

Section B: Questionnaire

5.	System Usability Scale (SUS) <i>*After using the website/application, please indicate your opinion accordingly.</i>	Strongly Disagree				Strongly Agree
		1	2	3	4	5
I think that I would like to use this application frequently		<input type="radio"/>				
I found the application unnecessarily complex		<input type="radio"/>				
I thought the application was easy to use		<input type="radio"/>				
I think that I would need the support of a technical person to be able to use this application		<input type="radio"/>				
I found that the various functions in this application were well integrated		<input type="radio"/>				
I thought that there was too much inconsistency in this application		<input type="radio"/>				
I would imagine that most people would learn to use this application very quickly		<input type="radio"/>				
I found the application very awkward to use		<input type="radio"/>				
I felt very confident using the application		<input type="radio"/>				
I needed to learn a lot of things before I could get going with this application		<input type="radio"/>				

6. Overall, I would rate the user-friendliness of this product as:



APPENDIX C

Original SUS Statements	The Modified SUS Statements	The SUS Statements used in this Study
I think that I would like to use this system frequently	I think that I would like to use this product frequently	I think that I would like to use this application frequently
I found the system unnecessarily complex	I found the product unnecessarily complex	I found the application unnecessarily complex
I thought the system was easy to use	I thought the product was easy to use	I thought the application was easy to use
I think that I would need the support of a technical person to be able to use this system	I think that I would need the support of a technical person to be able to use this product	I think that I would need the support of a technical person to be able to use this application
I found that the various functions in this system were well integrated	I found that the various functions in this product were well integrated	I found that the various functions in this application were well integrated
I thought that there was too much inconsistency in this system	I thought that there was too much inconsistency in this product	I thought that there was too much inconsistency in this application
I would imagine that most people would learn to use this system very quickly	I would imagine that most people would learn to use this product very quickly	I would imagine that most people would learn to use this application very quickly
I found the system very cumbersome to use	I found the product very awkward to use	I found the application very awkward to use
I felt very confident using the system	I felt very confident using the product	I felt very confident using the application
I needed to learn a lot of things before I could get going with this system	I needed to learn a lot of things before I could get going with this product	I needed to learn a lot of things before I could get going with this application

The original SUS statements [42], the modified SUS statements [41], the SUS statements used in this study.

APPENDIX D

Problems Severity
0: Not a problem
1: Cosmetic: need not be fixed unless extra time is available on project
2: Minor: fixing this should be given low priority
3: Major: important to fix, so should be given high priority
4: Catastrophic: imperative to fix this before product can be released

Relevance	Severity
The registration form has too many fields that the user does not find it necessary to fill up and most of the fields are not marked with (*) meaning is not important to fill (Home telephone number).	2
The "Select issue type" in "Contact us" form is complicated and has many details that confuse the user.	2
After clicking "Contact us" an intrusive popup message showed up "SOUQ: We are contacting Souq.com....." which confuses the user.	3
Different types of mobile phones in the same classification (LG & Apple).	4

Data entry	Severity
Text editing not easy in changing the name in the registration form, using the keyboard and small screen.	3
Inserting the card details is not easy for conducting the payment as the payment card has long number and the buttons are small to click quickly.	3
Very difficult to fill the forms fields as there is no Next/Tap button to move on to the next field.	4
Small area to click, the user presses the "name" text field few times to get the cursor inside.	1
Changing the username was difficult (e.g. Select the text and deleting).	1
Choosing the birthday was difficult and not easy to fill, small area to click the numbers.	3

Low input solution, it's hard to operate GUI widgets without a mouse, and it takes longer and are more error prone. Test entry is particularly slow and littered with typos.	3
--	---

Design/Layout	Severity
Choosing gender was confusing because the menu appeared from the bottom and not from where the user press the button.	3
The yes button in the pop up message after choosing log out is confusing as it located in the left where it should be in the right, as the language is Arabic.	1
Doesn't Automatically go to the main page after the user finished registration.	3
The registration button has no eye-catching colour or design so the user can easily locate and press.	1
The message "Move the page from the edges to browse other offers" confuses the user (intrusive popup pages) while looking for a product of mobile phone.	2
The buttons "Buy now" confuses the user when trying to add an item to the basket as the name does not reflect the function.	2
The product information/description and specifications can't be browsed easily as their links is not clearly presented for the user.	2
The user finds the content display/Navigation in the main page is not easy to browse through for the user to find what he/she is looking for.	2
The user can't easily locate the author name in the author list after putting all the specification via the filter button.	4
The user can't find the authors name in the author section as there are too many author names.	4
Using the filter is difficult and confuses the user because there are a lots of options.	3
The keyboard window cover the Email text field while filling in the registration form.	2
The user can't easily find the Basket button after he/she find the product.	3
The user can't easily add a product to the Basket as there is no "add to the basket" button.	3
The authors' names list contains some names written in English and some others written in Arabic which really confuses me.	2
The filter button has only a symbol and has no a name to reflect the function which usually lead the user to miss it.	2
The cancel button is written in English after typing in the search bar, although the chosen language was Arabic.	2
The cancel button placed in the right position (where it should be place in the right, Arabic format).	1
Cancel button remove the whole typed texted completely.	1
The results were not sorted by price after using the search function.	3
When writing in the search bar, the writing direction is towards the right, where it should be toward the left as Arabic language format.	3
The button "Slide to pay securely" Should be written in Arabic in the payment page.	3
The "Processing ..." page is written in English after entering the payment details.	3
The "Payment Confirmation" page is written in English after the "processing" page.	4
The "Continue" button is written in English in the "Payment Confirmation" page.	2
The final "Payment Confirmation" page is written in English.	4
Button for increase and decrease the amount doesn't handle the '0' amount.	3
Choosing the city is confusing while filling a new address as the menu appears from the bottom not from where the button was pressed.	3
Intrusive popup pages "Souq... An error has occurred. Please try again later in few minutes..." while trying to add an address confuses the user.	2
After filling all the payment information the user lost all the entered data by mistakenly visiting another page.	2
The "the verified by VISA" page was written in English.	1
Choosing the district was not easy (Scroll down a long list to find the user's district).	2
The user can't delete more than one item at a time in the basket page.	2
Putting the Card dates was confusing as the drop menu showed up from the bottom.	2
The option to make the address as a main address is confusing as there is no labels or text to explain whether it is "On" or "OFF" in add new address page.	2
The message about the house number is confusing as it says the "house number" in English where the rest of the message is Arabic.	2
The user can't buy only one product from basket.	2
The number in challenge question is written in English in "contact us" page.	3

The Challenge question doesn't accept Arabic numbers where the question is written in Arabic.	2
The user can't copy the "Order ID" from "Track your order" page.	2
The user Lost all the data entered in "Contact us" after getting out the page.	2
The Challenge question doesn't accept the answer while the keyboard is Arabic.	3
The user find it difficult to find the colour for the product mobile phone (Galaxy) as it's not clearly stated or presented.	3
The user couldn't easily find the description of a mobile product (Galaxy) after entering a product page.	3
The user can't locate the filter button as its button is small and its place not clear.	3
Using the filter options are complicated as the user needs to press "Apply" for each specification.	2

Completeness	Severity
Difficult to choose the username, there is no guidance or suggestions especially when the username is already taken by another user.	3
The system doesn't tell me what was wrong with my username and I have to change and guess the problem myself (e.g. does not accept the symbol "&").	2
After the user finishes the registration, no activation code was sent to the email, in this case anyone would register on behalf the user using his/her email.	3
The user can't find the previous account name, in case the username was already registered before.	2
The popup message "WRONG_VALUE_OF:lastname" does not give enough information about the mistake and how to solve it.	3
There is no enough information in the product list when looking for a mobile phone (Galaxy) (e.g. colour).	2
The user can't find the necessary information in the thumbnails view list to find the product he/she looking for.	2
The user finds the presentation of the products has not been effectively used the area of small screen of the smart phone (e.g. to present as much as possible).	2
There is no ability to search within the category of Mobile devices.	2
There is no advance search option so the user can search for a particular mobile phone with specified characteristics.	3
There is no filter option after the user used the searches function for a Samsung mobile.	2
The user can't use the filter after the search function in the main page.	2
The results were not sorted by price after using the search function in the main page.	2
There is no option to search mobile by colours or storage sizes.	2
The system doesn't show similar items while browsing an item.	2
Can't use a search function inside the book category.	3
Can't use a search function inside the publisher section.	2
After adding the address there is no confirmation message or email to the user.	4
The label text for Landline was missing and not clear for the user.	2
Choosing the district in the registration form confuses the user as there is no classification by area or county and there is no search function.	3
The "Payment Options" page gives only one choice "The user account" and cannot choose to pay by card or anything else.	4
The Challenge question is confusing the user as no justification/description for its purpose.	2
The user didn't get a feedback after sending the request through "contact us" form.	3
There is no Date for the Order ID in the page of "Track your Order".	3
Can't use a search function inside the Author section.	2

Visibility	Severity
Registration option/button was difficult to find from the main page of the application (e.g. the user has to visit many pages to find it).	3
Can't find the Mobile section easily from the homepage and had to search different pages to find it.	3
Can't find the filter option/button inside the "Samsung" category/section in order to set the mobile specifications (e.g. model, colour and storage size).	3
Can't locate the Book sections from the main page of the application and from the categories page.	2
Can't find the filter option/button inside the "Book" category/section in order to choose the book specifications (e.g. author name, publisher and the type).	3
Can't easily Locate the Filter option/button from the main page.	3
Can't find the "Order ID" for the purchased product, in order to track the product or to send a request	4

about it.	
Can't easily find the "Contact us" option/button from the main page, in order to contact the support team.	3
Can't find the correct request type in the "Select issue type" in the "Contact us" page form.	2

Download delays	Severity
Slow loading time after pressing the registration button in the registration page.	3
Slow loading time when the user open pages through the application (especially Main & Basket Page).	3
Adding the item to the basket took very long.	3
Slow page Loading time when trying to visit the page "Review Order/مراجعة الطلب".	3
Slow page Loading time after clicking the '+' to increase the quantity of the product.	3
Slow page Loading time after clicking the "The user account" in the payment page.	3
Slow loading for The "Payment Confirmation" page.	3
Slow page loading when pressing the button "more".	3

Correctness	Severity
The Basket named as "Shopping" which confuses the user MP.	2
The system doesn't allow the user to register and show a message "The server is not working".	2
Although the user choose the sup-category Samsung, an apple phones showed in the list.	3
The search engine fails to provide relevant information ("inaccurate results") in the homepage of the application when trying to search for a product (e.g. football).	3
The user can't find the publisher, although the name was entered مكتبة رير رير.	2
Deleting button was labelled as editing in the basket page which usually mislead the user.	3
Phone number in the "Add New Address" doesn't accept Numbers.	4
Landline number in the "Add New Address" doesn't accept Numbers and there is no guidance.	3
The first try to add an address was un successful and the system did not give any feedback.	1
The user directed to a webpage site version instead of the application after the payment went through.	3
There is no way to go back to the application after the payment was proceed successfully.	2
The "Payment Options" page doesn't response to the user (is not responsive) and all buttons cannot be clicked.	3
The application crashes after choosing the payment choice "The user account" (Three times crashes).	4
The application crashes after trying to delete an item from the basket.	4
The application crashes twice for no reason while adding a new address.	4
The delete button didn't work in the basket page.	3
The button "Add address" is not responding (none responsive) and cannot add my address.	3
Unresponsive "Card number" field while adding the payment card details.	4
The button "Edit" is not responding while the user tries to press it many times.	3
The "Order ID" text accept any fake number in the "contact us" form.	4
The application crashes while the user was trying to look for the "Order ID".	4
The filter doesn't give relevant results in the book section.	3
The home button is not responding in the main page.	3

Comprehension	Severity
The registration button is not clear for the user whether it's for a new registration or not.	2
The user does not know which name should he/she uses, the first or second for the author name.	2
The names "رير" and "مكتبة رير" confuses the user as they look very similar.	2
The further information button "!" confuses the user of what it actually does.	2
The Arabic label text for Mobile was "هاتف" (Phone) is not clear for the user (e.g. mobile phone or telephone).	1
The button "Slide to pay securely" is confusing the user as the user has to slide the button instead of pressing.	2
Duplicate page titles confuses the user "Payment Confirmation" while paying for a product.	1
Filling the phone number confuses the user as he/she does not know to put a zero or not.	1
The fields City and Area are confusing as the user Could not differentiate between them.	1
The button Add address is confusing the user (e.g. is it for a new address or for a second address?).	1
The button "Revise the order" confuse the user as it doesn't mean for paying.	2
There are two buttons "Contact us" and that's confuses the user.	3

A message "The server is not working" confuses the user as he/she does not know what to do. The message "we are contacting Souq.com. Thanks for your patience." after clicking "Contact us" confuses the user as there is no point of it for the user.	3 2
---	--------

APPENDIX E

Usability Test Observation Sheet

Date:.....

Participant's Identification Number:.....

Task#

- Time: ▪ Task allocated time:

Session Starts at:.....

Session ends at:

No	Usability Problems Discovered
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

	<ul style="list-style-type: none">..............................
--	---

Task Completion Rate:

Successful **Unsuccessful**

Notes:

.....
.....
.....
.....
.....