

2008

Developing Measures of Wireless Game Quality: A Three-Step Approach

Clarry Shchiglik

Victoria University of Wellington, klarik@xtra.co.nz

Stuart J. Barnes

University of East Anglia, Stuart.barnes@uea.ac.uk

Eusebio Scornavacca

School of Information Management Victoria University of Wellington Wellington, New Zealand, escornavacca@ubalt.edu

Follow this and additional works at: <http://aisel.aisnet.org/ecis2008>

Recommended Citation

Shchiglik, Clarry; Barnes, Stuart J.; and Scornavacca, Eusebio, "Developing Measures of Wireless Game Quality: A Three-Step Approach" (2008). *ECIS 2008 Proceedings*. 198.

<http://aisel.aisnet.org/ecis2008/198>

This material is brought to you by the European Conference on Information Systems (ECIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ECIS 2008 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

DEVELOPING MEASURES OF WIRELESS GAME QUALITY: A THREE-STEP APPROACH

Shchiglik, Clarry , School of Information Management, Victoria University of Wellington,
PO Box 600, Wellington, New Zealand, klarik@xtra.co.nz

Barnes, Stuart J., Norwich Business School, University of East Anglia, Norwich NR4 7TJ,
United Kingdom, stuart.barnes@uea.ac.uk

Scornavacca, Eusebio, School of Information Management, Victoria University of
Wellington, PO Box 600, Wellington, New Zealand, eusebio.scornavacca@vuw.ac.nz

Abstract

The potential market for m-commerce now consists of 2.5 billion users worldwide. In addition, as with more traditional computer-mediated environments, games are once again proving to be a dominant application. In particular, games are predicted to become the main driver in the m-commerce entertainment space. A key ingredient to the success of these wireless games is destined to be their quality. This is an area that has thus far been neglected by literature. The research performed in this paper develops, validates and demonstrates the usefulness of an instrument that specifically evaluates the quality of wireless games. Data collected by the instrument is grounded in subjective impressions. However, this data can be applied in quantitative analysis for the production of wireless game metrics such as the Wireless Game Quality Index. These objectives are carried out through a three-step methodology. First, an instrument to measure the construct of interest, wireless game quality is developed. Then the instrument's validity is examined and usefulness demonstrated in two phases of testing on wireless games. A five dimension structure of wireless game quality is revealed along with a validated 23-item instrument. The paper rounds off with conclusions and future directions for research and practice in this area.

Keywords: m-commerce, mobile; wireless; games, quality; evaluation; assessment; metrics.

1 INTRODUCTION

Games have been played since early civilisation (Shaw, 2004). They have become a pervasive component of culture and are used to serve a wide variety of purposes including entertainment, education and the refinement of skills (Crawford, 1982). The influence of computers to games has been particularly dramatic. The introduction of computers to games approximately 25 years ago has broadened the way in which games can be played and inspired the emergence of a growing industry. In 2007, the global market size of the computer games industry was measured at US\$37.5 billion and it is forecasted to increase in size to US\$48.9 billion by 2011 (PriceWaterhouseCoopers, 2007).

The emergence of wireless devices has now introduced a new dimension to games and promises greater market scope. In the space of just a decade, mobile phones have become one of the fastest adopted consumer products of all time (Chen, 2000; de Haan, 2000). Their explosive growth has led to over 2.5 billion subscribers worldwide (Wireless Intelligence, 2006). Moreover, the mobility and network capability offered by wireless devices opens a raft of capabilities beyond boundaries of wired connectivity (Siau and Shen, 2003). Already the growth of games played over mobile phones (referred to as wireless games from here onwards) can already be seen in a number of key markets worldwide, particularly in Asia and Europe (PriceWaterhouseCoopers, 2007).

Experience in the gaming market has shown that while a game's brand may initially be able to attract consumers, it will not guarantee its success. Rather, the long-run success of a game is largely determined by its quality as perceived by consumers (Kangas, 2003). In addition, games delivered over mobile communication devices operate in a different paradigm to those of other technologies, as dictated by differences in infrastructure and user behaviour. Therefore, it is interesting to observe that a good part of the problem with the initial wave of unsuccessful wireless games was due to a lack of understanding of quality from the perspective of consumers and how their needs and expectations could be met over the mobile medium.

Consequently, just as with any other mobile service, it is fundamental to have an appreciation and ability to measure customer perceived quality in order to achieve successful deployment of these types of games (Barnes, 2002; Kangas, 2003). Advancement in the understanding of customer perceived wireless game quality, and the ability to evaluate such for any given wireless game, will provide an increased understanding of customer's adoption behaviour and an ability to assess the likely success of commercialisation.

Accordingly, there exist several studies dedicated to understanding m-commerce quality (Barnes et al., 2001; Chae and Kim, 2001; Landor 2003). These research efforts represent a sound beginning for the study of quality in the m-commerce domain. However, they do not strictly focus on the intricacies of wireless games. By focusing on wireless games, distinctive features that would otherwise be unobserved through broad m-commerce research can be exposed. Moreover, only two of these studies have gone beyond initial scoping and delimiting of the field and specifically sought to establish an instrument for the evaluation of customer perceived m-commerce quality (Barnes et al., 2001; Chae and Kim, 2001).

The measurement of wireless game quality is similarly deficient within computer game. Three research efforts were identified as featuring game quality instruments (Desurvire et al., 2004; Gao, 2004; Schaefer et al., 2002). However, in addition to their scarcity, a common element amongst these studies is a lack of disclosure of the actual instrument featured in the study. Moreover, while the instruments are grounded within literature, no testing is performed to assert their validity for the measurement of game quality. Thus, it appears to date, academic literature has never disclosed a game-related quality instrument. Additionally, the identified game quality instruments are not specifically designed for wireless games. Yet, the various mediums over which games can be played possess distinct qualities (Crawford, 1982).

Given the importance for research into the area of wireless game quality and, at the same time, the deficiency of literature available at present in this field, the objective of the current study is to develop a customer perceived instrument for the measurement of wireless game quality and then to demonstrate the validity and usefulness of this instrument.

The remainder of the paper is structured as follows. The next section provides an overview of the research approach and design. This is followed by three sections that detail the results of the research within the three research steps. Note that, by design, details on the literature are integrated within appropriate locations in the paper's structure. The paper concludes with a discussion of the key research findings, limitations, and suggestions for further research and practice.

2 RESEARCH APPROACH

A number of researchers have studied and proposed procedural models for the development of instruments with better measures (e.g. Bagozzi, 1980; Churchill, 1979; Straub, 1989). Over time, these suggested procedures have grown to become commonly applied in the Information Systems field (e.g., Doll and Torkzadeh, 1988; Aladwani and Palvia 2002). Aladwani and Palvia (2002) have identified three generic common steps amongst many of these of procedural models: 1. conceptualisation; 2. design; and 3. normalisation.

The research framework of the current study, shown in Figure 1, features two phases that summarize the three generic steps of instrument development identified by Aladwani and Palvia (2002). The first phase involved qualitative methods and comprised of the conceptualisation step. This step aims to develop a candidate list of items based on a specified construct of interest and thus content validity is of key concern. Content validity was ensured by adopting a thorough approach involving three processes: literature review, focus groups and expert review. Conceptualisation commenced by referring to literature to expressly specify the wireless game quality construct. Then with consideration of the defined construct, data collected from four focus groups was analysed and combined with literature to produce an initial list of candidate items. This initial list of candidate items was then reviewed by experts.

The second phase involved quantitative methods and involved two steps, design followed by normalisation. The purpose of the design step is to refine the list of candidate items put forward from the previous step and derive an initial instrument. This involves instrument design and performing a series of analytical procedures on data collected from pilot tests and an explanatory survey. Construct validity and reliability are considered throughout the design step. Normalisation then sought to verify the instrument's validity and demonstrate it's usefulness by establishing norms. These procedures were carried out by performing analytical procedures on data collected from a subsequent explanatory survey.

3 STEP I: CONCEPTUALISATION

To begin with, the domain of the construct of wireless game quality was specified. This was carried out in consultation with literature. For the purposes of this research, wireless game quality is defined as a customer's evaluation of a wireless game satisfying their expectations. Previous attempts to define wireless game quality were not found. As a result, this definition is essentially a merger between related concepts within the fields of wireless games and quality (e.g. Cronin and Taylor, 1992; Hoyer and Hoyer, 2001; Oliver, 1980; Oliver, 1997; Pelkonen, 2004).

The next step of conceptualisation was to generate items that capture the domain of the construct of interest, wireless game quality. In order to ensure content validity it is essential the generated items are based on the specified domain (Churchill, 1979). Procedures that are beneficial for ensuring content

validity are a review of literature and exploratory data collection, both of which were undertaken in this study (Malhotra and Grover, 1998).

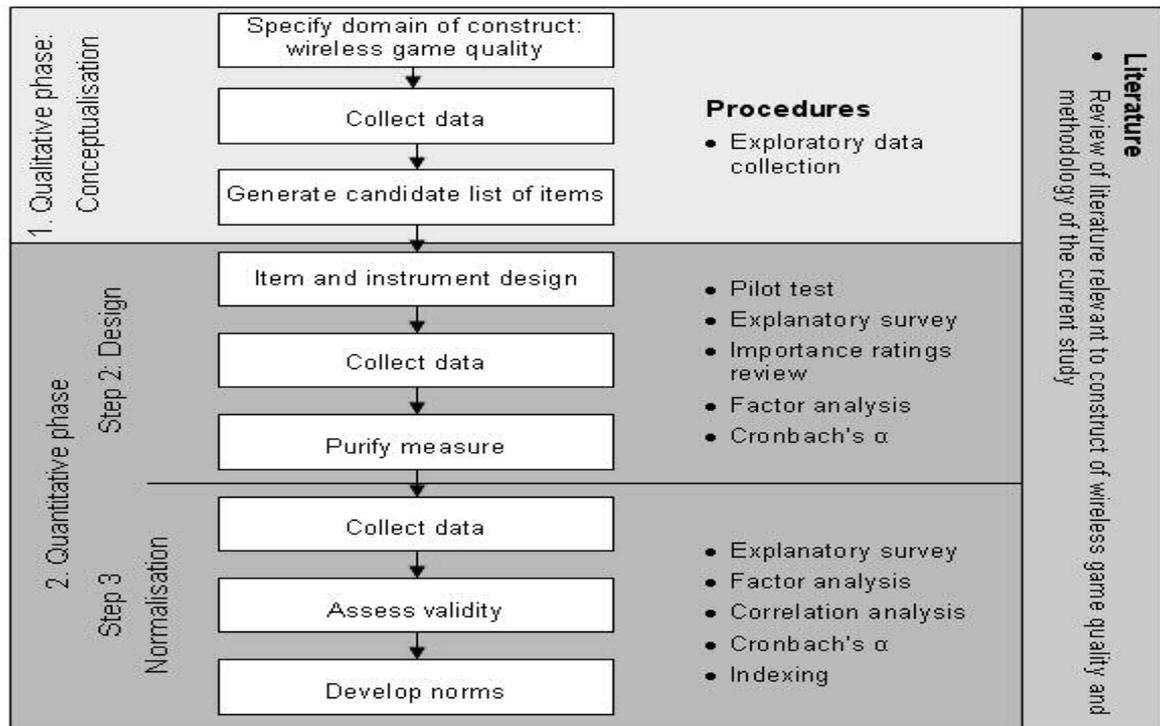


Figure 1 Research framework

A wide range of sources were reviewed in order to find literature relating to wireless game quality. This included journals, conference proceedings and postings on online community websites that concentrate on closely related areas. In addition, two types of exploratory data collection were performed, focus groups and an expert review.

Four focus groups were carried out, each consisting of six participants. Focus groups ran for between 90 – 120 minutes and featured three stages: (1) playing WAP games; (2) focus group discussion; and (3) a quality workshop. Data collected from focus groups were analysed using grounded theory techniques. In addition, focus groups offered the overall research a by-product of insight into raw customer perceptions towards wireless games. For greater detail behind the methodology and results and analysis that relate to these observations please refer to our previous publication (Shchiglik, Barnes and Scornavacca, 2005). The core focus of this paper is to emphasise how focus groups contributed to instrument development. Literature was combined with focus group findings in order to produce a total of 40 items conceptualising wireless game quality.

The initial list of 40 candidate items was then sent out to 12 experts for feedback to ensure the wireless game quality construct was adequately captured. Experts were either in the category of academic or industry, and were involved in either game research or development. Their knowledge offered an additional layer of insight into wireless game quality and, as a result, further enhanced content validity. From the 12 experts contacted, only eight responses were received. Three were from academia and five from industry. Once all responses had been collected, expert's comments were reviewed by researchers involved in the current study. This process resulted in seven changes to the instrument. Changes comprised of the addition of three items, the amendment of one item and the removal of three items.

4 STEP II: DESIGN

In practice, it is not feasible to administer all possible items that could potentially be used to measure wireless game quality nor the entire list of candidate items put forward from the conceptualisation step (Churchill, 1979). For this reason, it is necessary to purify the candidate list of items put forward from the conceptualisation step in order to retain a sample of items that represent the construct of wireless game quality and offer a highly correlated measure of the true score of wireless game quality. In addition, it is possible an initial instrument or items generated from conceptualisation may present respondents with operational issues. These operational issues are identified and resolved and purification is achieved in the current study with the application of analytical procedures of data collected from pilot testing and an explanatory survey.

Before pilot tests commenced, which customer perspectives would be measured needed to be addressed. This question arises as service quality has been previously operationalised with several different measurement approaches. Literature has sought to assess service quality with measurement of customer perceptions (i.e., service quality as perceived by a customer), customer expectations (i.e., expectations of the service quality that will be delivered), and importance ratings (i.e., the importance of a quality to a customer).

The approach applied to the current study is to measure customer perceptions and importance ratings. Customer perceptions consider the degree to which a quality has been attained, while importance ratings consider the importance of a particular quality. This approach is believed to capture the most important aspects of quality and at the same time keep the number of assessments that need to be made to a manageable number (Barnes and Vidgen, 2000). Furthermore, Mazis et al. (1975) suggest the combined measurement of customer perceptions and importance ratings is the most efficient approach to use where the objective is to predict behavioural intention or actual behaviour.

A pilot test was carried out to identify issues with the instrument and overall survey procedure. Eleven undergraduate students participated in the pilot test. Each participant was asked to play a WAP game for a period of time that they felt was sufficient to be able to then assess its quality. Feedback was received from participants through a questionnaire, which included the developed instrument, and also a one-to-one face-to-face interview.

Pilot tests revealed detailed insights into the instrument and survey process. Participants felt there were an excessive number of items and several items had either interpretation difficulty or questionable applicability to the assessment of wireless game quality. Feedback received from the pilot tests resulted in the rephrasing of three items and removal of five items. Thus, the instrument was reduced to 35 items.

An explanatory survey was performed in order to allow the instrument to be refined. The explanatory survey was carried out by evaluating three WAP games as they were the most common wireless game format available at the time and students were targeted as respondents as they are known to be a key market for wireless games. To increase the response rate, respondents were asked to evaluate only one game and were offered the opportunity to win one of two \$50 vouchers. This resulted in 86 responses, 85 of which were usable.

The demographics of the sample were as follows. Two-thirds of respondents were male and one-third female. The majority of the sample lay in the age brackets 16-20 (32%) and 21-25 (52%). Also, an overwhelming majority of the sample, 82%, had previously played a wireless game, but only about a third had played a WAP game.

Data collected relating to importance ratings offered the first means by which the instrument could be refined. As importance ratings measure the relevance of an item to the assessment of wireless game quality, they allow the instrument to be restricted to contain only salient items. Thus, one approach employed to ensure the final instrument contained items considered most relevant for the measurement of wireless game quality was the removal of items with a low importance rating. This was carried out

in a systematic process whereby items that fell into a category with the significantly lowest importance rating means were removed. This resulted in the removal of four items.

Factor analysis was then applied to the collected data to reveal the underlying factors of wireless game quality. Items removed due to low importance ratings were included in this analysis to help identify these underlying factors. Factor analysis was carried out based on two key criteria. Firstly, retain items that loaded greater than 0.55 on a single factor and less than 0.5 on all other factors. Secondly, retain factors that were found to consist of at least two items.

Factor loadings represent correlations between original item scores and factors. Thus, convergent validity is claimed when instrument items load highly on relevant factors. Conversely, discriminant validity is claimed when items load more highly on one factor than on others. Factor analysis also reveals sub-constructs of the overall construct of interest.

The Varimax factor rotation converged in 15 iterations and revealed eight factors. Three of these eight factors contained only one item. Such factors are common for factor analysis procedures performed at an early stage of an analytical process and the items within these factors are not considered to have a significant relationship with the focal construct of interest, i.e., wireless game quality (Churchill, 1979). For this reason, lone item factors were eliminated along with the items they represented. Furthermore, seven items were found to not have a factor loading over the 0.55 cut-off for any factor and were removed from the instrument. However, one of these items had already been removed due to low importance rating, thus, factor analysis resulted in reducing the instrument by nine items.

Factor analysis revealed five dimensions of wireless game quality that are distinct and uni-dimensional and thus, indicate evidence of both convergent and discriminant validity. These factors were: ease of use, ease of access to quality information, information response, aesthetic appeal and gaming experience. Explanations of the dimensions of wireless game quality are as follows:

- *Ease of use*: the ease of learning, playing and understanding the objective of a game. The dimension also considers ease with which it is possible to navigate.
- *Ease of access to quality content*: the straightforwardness with which accurate, relevant and easy to understand content can be accessed.
- *Information response*: the period of time taken to receive a response and the timeliness of information received.
- *Aesthetic appeal*: the attraction of a game base on the image it conveys and its use of multimedia.
- *Gaming experience*: qualities concerning flow and enjoyment from playing games, e.g., keeping a user's attention focused, the incentive of playing and the interest gained from the experience.

Reliability was the final analytical procedure of the design step. Cronbach's α was 0.9340 for the overall instrument and ranged from 0.7188 to 0.9194 for measurable dimensions. All α coefficients are in the range of acceptability (Nunnally, 1967).

Factor analysis revealed five underlying dimensions of wireless game quality. However, the dimensions of information response and aesthetic appeal were believed to contain an insufficient number of items to capture their intended meaning. Moreover, it is most important for each factor to be measured by a multiple number of items (Churchill, 1979). Information response contained two items and aesthetic appeal only one item. Hence, three items were added to the dimension aesthetic appeal and one item to the dimension information response.

As a whole, procedures undertaken in the design step demonstrated convergent and discriminant validity as well as reliability. The step also resulted in the removal of 18 items (pilot tests five items, importance ratings four items and factor analysis nine items), rephrasing of three items and subsequent addition of four items. Thus, the instrument was refined to contain 26 items.

5 STEP III: NORMALISATION

Normalisation was the third and final step of the research framework. A key focus of the step is construct validity. Normalisation involves a subsequent independent verification and validation of the instrument's construct validity as initially proposed previously in the design step. In addition, normalisation sets out to demonstrate the usefulness of the instrument by establishing norms. These procedures are achieved by analysing data from a second explanatory survey.

The second explanatory survey was carried out by evaluating three WAP games - Boy Racer, Riddler and Trivia Racer. Boy Racer is a multiplayer game designed to test memory, the purpose of Riddler is to solve word puzzles, and Trivia Racer is a multiplayer, multiple-choice general knowledge game where correct answers push the player along a racetrack. The survey used a five-point Likert scale where the anchors range from 1 ("strongly disagree") up to 5 ("strongly agree"). In addition, to encourage participation and the quality of response, a prize draw with a \$250 first prize and \$50 second prize was offered. These procedures resulted in 127 responses, of which 125 were usable.

Two-thirds of respondents were male and one-third female. The sample consisted primarily of respondents in the 21 to 25 (36%) and 26 to 30 (52%) age group bands. In addition, the majority of the sample (83%) had experienced playing a wireless game and a sizeable portion (43%) of the sample had played a WAP game. Importance ratings and perception scores were expressed in terms of their mean and standard deviation values.

The Varimax factor rotation converged in nine iterations and confirmed a five factor structure of wireless game quality, namely: ease of use, ease of access to quality information, information response, aesthetic appeal and gaming experience. The results are shown in Table 1.

No.	Item Description	Factor				
		Gaming experience	Ease of access to quality info.	Ease of use	Information response	Aesthetic appeal
1	Easy to use			0.7917		
2	Easy to learn and operate			0.8196		
3	Knew what I had to do			0.8082		
4	Easy to navigate			0.6779		
5	It is easy to find things in the game			0.6282		
6	Provides accurate content		0.7324			
7	Provides relevant content		0.7974			
8	Provides easy to understand content		0.6244			
9	<i>Provides consistent navigation</i>				0.5294	
10	<i>Provides fast navigation to what I intend to find</i>				0.5685	
11	Provides timely content				0.6295	
12	Has a fast response time				0.8493	
13	Provides feedback quickly				0.8241	
14	Meets my expectations	0.5726				
15	Provides an enjoyable experience	0.6944				
16	Provides value for money (7c per click)	0.5621				
17	Keeps my attention focused	0.6311				
18	<i>Provides an appropriate level of challenge</i>	0.5215				
19	Provides an incentive to play	0.6472				
20	Benefits are gained from playing with others	0.6038				
21	Has innovative game play	0.7251				
22	Interesting to play	0.6292				
23	Has visually pleasing graphics					0.7630
24	Looks professionally designed					0.8338
25	Has an attractive appearance					0.8488
26	Design is suited for the mobile phone					0.6164

Table 1 Factor analysis (factor loadings 0.5 and above, second explanatory survey)

Items 9 and 18, both failed to produce a factor loading greater than 0.55 for any factor. Item 10, on the other hand, loaded above the 0.55 cut-off for a different factor than previously identified. In the earlier factor analysis, item 10 loaded 0.5520 for the factor ease of access to quality information, whereas in this factor analysis the loading was 0.5685 for information response. There is the possibility that the exploratory nature of the first factor analysis caused the item to be initially and erroneously placed into the ease of access to quality information factor. Further research is required to confirm this supposition. However, in any case, the suggestion of a lack of discriminant validity of this item was sufficient to warrant its removal from the instrument. Moreover, in the earlier factor analysis this item had already indicated a discriminant validity concern, but was retained as part of the ease of access to quality information factor as the discrepancy was only marginal. A closer inspection of the item's phrasing, "fast navigation to what I intend to find", further supports these empirical findings; "fast" is indicative of information response, while "navigation to what I intend to find" would likely relate to easy access to information.

The inability of item 18 to meet the 0.55 loading cut-off is particularly interesting as challenge is considered amongst flow literature to be an integral part of the interactive experience (Csikszentmihalyi, 1975). However, perhaps the term 'challenge' offers respondents an unintended interpretation that inspires connotations of work, effort and exercise – all of which are incompatible with the factor gaming experience, the factor which had earlier encompassed the item. With this premise, the removal of item 18 reduces the potential ambiguity of the instrument. Further explanation of the absence of item 18 from the gaming experience factor can be gained from considering the scope of item 22, "Interesting to play". The two items are in many ways related. Moreover, a review of the scope of item 22 reveals it encapsulates the intended meaning of item 18.

Flow literature asserts that if a game is not challenging it will become boring quickly; conversely, if it is too challenging a player will feel frustrated and discouraged (Mallon and Webb, 2000). Comments received from the exploratory data collection procedures of the conceptualisation step as well as the first and second explanatory surveys included: "Boring game, not challenging, or interesting and no graphics" (first explanatory survey), "Pretty straight forward game that does not capture people's interest. No challenge" (first explanatory survey) and "...the questions were too hard." (focus group discussion). The comments present a strong association between the concepts of challenge and whether a game is interesting. They also express respondent's dissatisfaction with games that are either overly easy or difficult. As a result, the inclusion of item 22, an item that determines whether a game is interesting to play, acts to effectively include the intended qualities of 'challenge' into the gaming experience factor.

The 23 items retained in the final	Number of items	Cronbach's α
Gaming experience	8	0.8881
Ease of use	5	0.8856
Ease of access to quality information	3	0.8327
Information response	3	0.8297
Aesthetic appeal	4	0.8303
Overall	23	0.9320

Table 2: Cronbach's α coefficients (second explanatory survey)

Convergent validity is concerned with the extent to which multiple measures of the same construct agree with each other (Campbell and Fiske, 1959). The correlation coefficient between the instrument and overall quality was found to be 0.771 with a p value of 0.000. Discriminant validity refers to the extent to which measures of different constructs are distinct (Campbell and Fiske, 1959). As we can see from Table 1, the factors are clear and distinct in their loadings, with no overlap.

The usefulness of the instrument was demonstrated by developing norms. A raw score obtained from an instrument measurement is on its own not particularly informative; however, establishing norms

created a benchmark and thus allowed meaningful comparisons and interpretations of the quality of the three evaluated wireless games. Norms were developed by calculating an index, referred to in the current study as the Wireless Game Quality Index (WGQI). The WGQI is calculated by dividing the weighted score by the maximum score. The weighted score represents each respondent's perception score multiplied by the importance rating attached to it and the maximum score represents the mean importance rating multiplied by 5 – the maximum rating attainable. Overall, Riddler is benchmarked highest with an overall WGQI of 0.72. Riddler is relatively closely followed by Trivia Racer, which has a WGQI of 0.66 and Boy Racer is well below the other two games with a WGQI of 0.55.

Further, more detailed, comparisons between the three wireless games can be made by applying the WGQI with respect to the five dimensions of wireless game quality, namely, ease of use, ease of access to quality information, information response, aesthetic appeal and gaming experience. The required calculations commenced by grouping data according to dimension item groupings. Then the total score for each dimension was divided by its maximum score to derive a WGQI. The results of these calculations are shown as a graphical representation in Figure 2. These comparisons are particularly beneficial as they permit a closer examination into why some games fared better than others in terms of the overall WGQI. Note the scale have been restricted to 0.45 to 0.85 to help draw clearer comparisons.

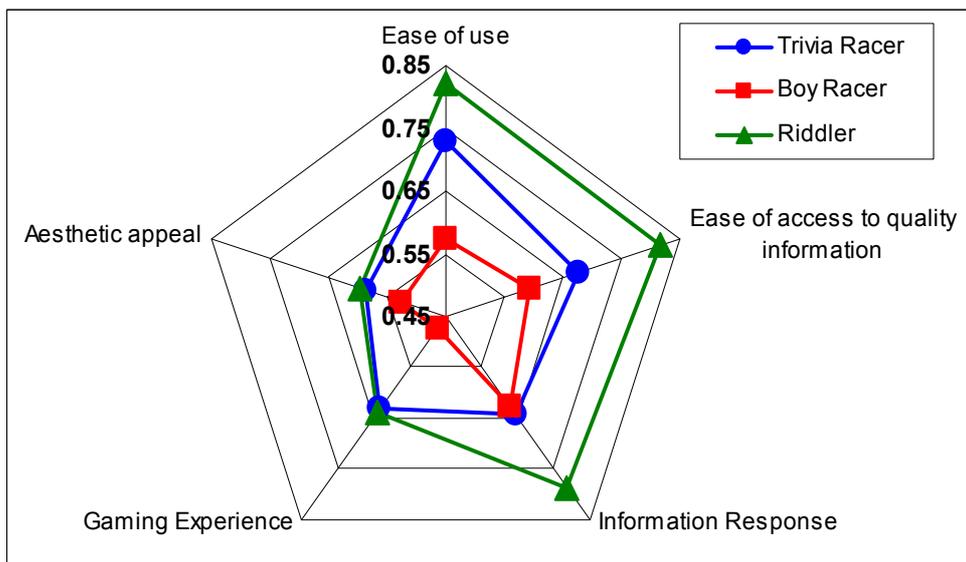


Figure 2: Radar chart of WGQI dimensions with respect to the three evaluated games

Overall, Riddler was found to have a substantially higher WGQI score than the other two games with respect to the dimensions of ease of use, ease of access to quality information and information response. However, in terms of the other two dimensions Riddler scored similarly to Trivia Racer. In addition, it was found that while Trivia Racer scored substantially higher than Boy Racer at an overall level, the two games had a similar WGQI in regards to the information response dimension.

The relatively high WGQI score achieved by Riddler for the ease of use dimension reflects the game's uncomplicated nature. Riddler simply requires players to read a riddle and then click to see the answer. There is no score keeping or game mechanism to determine whether the player was correct. Conversely, Trivia Racer, for example, presents a more complicated task proposition. For example, a game cycle for Trivia Racer requires typically upwards of a dozen screens to be downloaded and viewed compared to only two for Riddler. Thus, by its very nature, Trivia Racer is expected to be more difficult to use.

A plausible explanation for the similarity in WGQI scores between Trivia Racer and Boy Racer for the information response dimension rests with their game design. Both games are multiplayer and are thus

heavily reliant upon network connectivity. Moreover, the way in which multiplayer gaming is implemented in both games is identical: players are pitted against one another to complete a race course and progress by correctly answering questions. They answer questions synchronously and independently of each other, thus, their ability to compete against one another is heavily reliant upon the quality of their network connection, i.e., should a player receive favourable connectivity then they will be in an advantageous position.

Compared to Trivia Racer and Boy Racer, Riddler scores substantially higher for the information response dimension. In terms of information response qualities, Riddler is most notably distinctive for being single player and also excluding graphics other than what is seen on its main introduction screen (i.e., less data needs to be downloaded). As a result, Riddler is comparably less reliant upon network connectivity. Results regarding responsiveness find multiplayer games disadvantaged due to their reliance upon mobile network connectivity. This reaffirms the need for game design to cater for limitations of mobile networks.

Another interesting observation regarding WGQI results concerned the aesthetic appeal dimension. This was the only dimension in which all three games score below 0.6 on the WGQI scale. It is believed that WGQI scores for this dimension are a reflection of the restricted capabilities of the format of the wireless games evaluated – i.e., WAP – to cater for interactive multimedia capabilities that are typically suited for games such as sound and moving objects. Moreover, aesthetic appeal was noticeable for having the lowest spread of WGQI scores. For this dimension, WGQI scores ranged from 0.53 to 0.60. It is possible that the three evaluated games bear many similarities in terms of aesthetic appeal. Alternately, it is also possible that the limited multimedia capability of WAP presents little opportunity for games to differentiate themselves in terms of this dimension.

6 CONCLUSIONS

This paper aimed to develop and demonstrate the usefulness of a valid customer perceived instrument that specifically caters for the measurement of wireless game quality. Research was carried out in the form of a three-step, two-phase investigation. The first phase focused on the development of an instrument that conceptualised the construct of wireless game quality. The second phase focused on validity testing and demonstrated the usefulness of the developed instrument through a process design and then normalisation.

Results of the two-phased investigation uncovered five dimensions of wireless game quality: ease of use; ease of access to quality information; information response; aesthetic appeal; and, gaming experience. The results provided evidence for the psychometric properties of the 23-item instrument to measure wireless game quality. A rigorous process to ensure content validity and tests for convergent and discriminant validity as well as reliability demonstrated evidence of a robust instrument.

Results from the study indicate the developed instrument is a useful diagnostic tool for the assessment of wireless game quality from a customer's perspective. The instrument derives a WGQI score, which can be applied at an overall wireless game quality level by using the entire 23-item instrument or at a specific wireless game quality dimension level, i.e., using a sub-scale of one of the five dimensions of wireless game quality. Furthermore, the instrument can be applied either as a cross-sectional survey to provide a benchmark against other wireless games or longitudinally to ascertain the consequences of amendments and the affect of time. For industry, the developed instrument serves as a guide to realise customer needs and wants and accordingly produce wireless games that effectively meet these requirements.

The current study presents benefits to both academia and industry. The study extends the current state of knowledge of m-commerce, games, quality and respective subfields of wireless games, m-commerce quality and game quality. More specifically, as wireless game quality has not been previously studied and documented in the literature, the current study also marks the commencement of research in the specific field of wireless game quality. Furthermore, the instrument developed in the

current study represents a direct extension of existing instruments from m-commerce and game quality research. In much the same way, the instrument developed here may be modified for other domains, whether they be other specialised service areas such as m-commerce auctions or technology mediums such as dedicated mobile gaming devices.

Key limitations of the research performed concerned focus groups and explanatory surveys. Both procedures exposed respondents to a restricted perspective of wireless games. Focus groups consisted only of WAP games. In addition, both the first and second explanatory surveys' respondents experienced varying degrees of the evaluated wireless games before making their assessments. While some respondents had a winning experience others had a losing experience. Respondent's perceptions of a game are likely to have been affected by their success. Moreover, in cases where respondents had lost, they may have only experienced a partial representation of a game, for example, not seeing the benefits that a game presents upon winning.

There are number of possibilities for future research to advance the progress made in the current study. This is particularly so with respect to the refinement and improvement of the developed instrument as "no good canvas is completed in a first attempt" (Pitt et al. 1997, p. 218). The current study presents a first step in the development of a valid instrument that specifically caters for the measurement of wireless game quality and there remains scope for further improvement. A possibility for future research is to conduct further exploratory data collection for the purpose of understanding consumer perceptions of wireless game quality with respect to wireless game formats other than WAP, namely, message-based, downloadable games and other available formats.

The wireless game qualities identified in the 23-item instrument can also be extended to understand key relationships between wireless game qualities and usability heuristics (Haag et al., 1996). For example, the heuristic "does the interface include scorekeeping" would intuitively be expected to have a strong link back to the quality "provides relevant content". Matrices of qualities and heuristics need to be developed in order for key relationships to be uncovered and, as a result, the voice of the customer to be deployed throughout wireless game design and development. It is not necessary, nor likely to be feasible, to model all plausible relationships. Rather, the aim is to understand the strong relationships in order to provide wireless game developers with critical guidance. The translation of customer demanded qualities into specific heuristics in such a fashion offers great value as it is ultimately the customer that determines the quality and success of a wireless game.

References

- Aladwani, A.M. and Palvia, P.C. (2002). Developing and validating an instrument for measuring user-perceived web quality. *Information and Management*, 39 (6), 467-476.
- Bagozzi, R.P. (1980). *Causal models in marketing*. New York: John Wiley.
- Barnes, S. J. and Vidgen, R. T. (2000). Information and interaction quality: Evaluating Internet bookshop web sites with WebQual. In *Proceedings of the 13th International Bled Electronic Commerce Conference*. Bled, Slovenia.
- Barnes, S. J., Liu, K. and Vidgen, R. T. (2001). Evaluating WAP news sites: The WebQual/M approach. In *Proceedings of the 9th European Conference on Information Systems*. Bled, Slovenia.
- Barnes S. J. (2002). Provision of services via the Wireless Application Protocol: A strategic perspective. *Electronic Markets*, 12 (1), 14-21.
- Campbell, D.T. and Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Chae, M. and Kim, J. (2001). Information quality for mobile Internet services: A theoretical model with empirical validation. In *Proceedings of the Twenty-Second International Conference on Information Systems*. New Orleans, U.S.A..
- Chen, P. (2000). Broadvision delivers new frontier for e-commerce. *M-Commerce*, October, 25.
- Churchill, G. A (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, 16 (1), 64-73.

- Crawford, C. (1982). The art of computer game design. Washington State University, Vancouver. Available from: <http://www.vancouver.wsu.edu/fac/peabody/game-book/Coverpage.html> (accessed 7 October 2006)
- Cronin, J. J. Jr. and Taylor, S. A. (1992). Measuring service quality: A re-examination and extension. *Journal of Marketing*, 56 (3), 55-68.
- Csikszentmihalyi, M. (1975). *Beyond Boredom and Anxiety*. Jossey-Bass, San Francisco, CA.
- de Haan, A. (2000). The Internet goes wireless. *EAI Journal*, April, 62-63.
- Desurvire, H., Caplan, M. and Toth, J. A. (2004). Using heuristics to evaluate the playability of games. In *Proceedings of the Conference on Human Factors in Computing Systems*. Vienna, Austria.
- Doll, W. J. and Torkzadeh, G. (1988). The measurement of end user computing satisfaction. *MIS Quarterly*, 12 (2), 259-274.
- Gao, Y. (2004). Appeal of online computer games: A user perspective. *The Electronic Library*, 22 (1), 74-78.
- Haag, S., Raja, M. and Schkade, L. (1996). Business transformation on the Internet. *European Management Journal*, 17 (5), 466-476.
- Hoyer, R.W. and Hoyer, B.Y. (2001). What is quality? *Quality Progress*, 7, 53-62.
- Kangas, S. (2003). *Benchmarking Literature Review*, European Commission, Brussels.
- Landor, P. (2003). Understanding the foundation of mobile content quality: A presentation of a new research field. In the proceedings of the 36th Hawaii International conference of System Sciences, 1-10. Waikoloa, Hawaii, U.S.A.
- Malhotra, M. K. and Grover, V. (1998). An assessment of survey research in POM: From construct to theory. *Journal of Operations Management*, 16 (4), 403-423.
- Mallon, B. and Webb, B. (2000). Structure, causality, visibility and interaction: propositions for evaluating engagement in narrative multimedia. *International Journal of Human-Computer Studies*, 53 (2), 269-287.
- Mazis, M. B., Olli, T. A. and Kleppel, R. E. (1975). A comparison of four multi-attribute models in the prediction of consumer attitudes. *Journal of Consumer Research*, 2, 38-52.
- Nunnally, J. C. (1967). *Psychometric theory*. New York, NY: McGraw-Hill.
- Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, 17 (November), 460-469.
- Oliver, R. L. (1997). *Satisfaction: A behavioral perspective on the consumer*. New York: Irwin/McGraw-Hill.
- Pelkonen, T. (2004). Anticipating content technology need. E-Content Reports. ACTeN paper.
- Pitt, L. F., Watson, R. T. and Kavan, C. B. (1995). Service quality: A measure of information systems effectiveness. *MIS Quarterly*, 19 (2), 173-187.
- PriceWaterhouseCoopers (2007). *Global Entertainment and Media Outlook: 2007-2011*. Available from: <http://www.pwc.com/> (accessed 23 November 2007).
- Schaefer, C., Enderes, T., Ritter, H. and Zitterbart, M. (2002). Subjective quality assessment for multiplayer real-time games. In the proceedings of the 1st Workshop on Network and System Support for Games, 74-78. Braunschweig, Germany.
- Shaw, I. (2004). *The Oxford History of Ancient Egypt*. U.S.A.: Oxford University Press.
- Shchiglik, C., Barnes, S. J. and Scornavacca, E. (2005). Customer perceptions towards mobile games delivered via the wireless application protocol. In *Unwired Business: Cases in Mobile Business*. London: IRM Press.
- Siau, K. I. and Shen, Z. (2003). Mobile communications and mobile services. *International Journal of Mobile Communications*, 1 (2), 3-14.
- Straub D. W. (1989). Validating instruments in MIS research. *MIS Quarterly*, 13 (2), 146-166.
- Wireless Intelligence. (2006). Worldwide cellular connections pass 2.5 billion. Available from: <http://www.ovum.com/go/content/c.377.66726> (accessed 12 November 2006)