MyWorld: Testing female preferred gaming criteria

Richard Cooke
Deparment of Computer Science
University of Cape Town
rjec@global.co.za

Jonathan Jedeikin
Deparment of Computer Science
University of Cape Town
jjedeiki@cs.uct.ac.za

Douglas Lepar
Deparment of Computer Science
University of Cape Town
dlepar@cs.uct.ac.za

ABSTRACT
In this paper we describe research conducted to determine the validity of ten criteria previous research has shown that females enjoy in computer games. These criteria were embodied in a graphical online multiplayer game, *MyWorld*, which was evaluated by users in order to determine the validity of the embodied criteria. The game itself was a novel idea in which players could add their own images to the virtual world. It was found that seven of the ten criteria are preferred by females and should be incorporated in gender aware games.

Categories and Subject Descriptors
H.1.2. [**User/Machine Systems**]: Human factors

H.5.1 [**Multimedia Information Systems**]: Artificial, augmented, and virtual realities.

K.8 [**Personal Computing**]: Games

General Terms
Measurement, Performance, Design, Experimentation, Human Factors.

Keywords
Gender, Chatrooms, Java, Client, Server, Database, Massively Multiplayer Online Games, Virtual Environments, User Participation.

1. INTRODUCTION
There is a male gender-bias towards almost all aspects of computers. This bias includes gender-biased social norms and values that discourage young girls from computers and cause a dominant majority of male computer science students, programmers and game designers. It stems largely from the scarcity of computer games that are enjoyable for young girls [5]. The discipline of computer science is therefore limited due to the lack of female participation.

1.1. Preferred Female Gaming Criteria
In her book, *Gender inclusive game design*, Graner-Ray clearly summarises and details research in psychology and marketing by game development firms. Her research is also supported by other academics. They include [6], [8], [9], [11] and [16]. Based on this literature, we created ten female preferred gaming criteria. These ten criteria aim to encompass all the results of the research referenced above. They are:

1. No over-sexualised avatars: Representations of the player in the game world can be attractive but their sexual characteristics are not exaggerated to an extent that is disliked by players.
2. Indirect competition: Players can compete against other players by doing their best, but not by affecting another player's performance. (E.g. running a race as opposed to a boxing match).
3. Collaboration: Players can act jointly with other players to achieve a common goal. (Note that a goal can be continuous. E.g. designing a theme or amassing points).
4. Emotionally involving: Players experience emotions when playing the game (e.g. they feel happy, sad, scared, disgusted or amused). These emotions should be caused by game elements (e.g. other players or their creations) and not by the interface.
5. Non-violent: No physical harm can be caused to any inanimate or animate object.
6. No player death or insurmountable penalties: A player cannot make a wrong decision that causes her to be removed from a game or have to restart.
7. No sexual prejudice: Players do not experience sexual discrimination.
8. Simple graphics: 2D graphics are adequate for enjoyment.
9. Creativity: Players can express their own individuality.
10. Process orientated: The enjoyment of the game stems from the actions players perform in the game continuously, rather than from winning the game.

1.2. Research Objectives
The above criteria were embodied in an online multiplayer game, called *MyWorld*, which we evaluated through user participation to determine the validity of the embodied criteria. The game itself was a novel idea in which players could add their own images to the virtual world.

Specifically, our research aimed to:

- Create criteria females prefer in computer games, based on a summary of past research,
- Create, with user participation, a computer game embodying these criteria,
- Test whether the criteria were embodied in the game,

- Determine females' attitudes towards the game and the criteria by means of a survey.

The results of this study would determine whether these criteria should be incorporated in computer games, to make future games gender inclusive. From a technical point of view, the research also tested the feasibility of creating a virtual world where users could add their own images to the game, as well as placing pre-drawn images the game had provided to them. To our knowledge, this has not been attempted before. For instance, in *The Sims Online*, players can place virtual furniture provided with the game, but are not able to add their own furniture designs. Additionally, if *MyWorld* were successfully created and enjoyed by young girls, it might be a useful educational tool for schoolchildren. Being both non-gender biased and non-violent, it could be suitable to teach young children collaboration and creativity, while accustoming them to using a computer. Our participants were all approximately fourteen years of age, to help us create a game that would encourage young girls to enjoy computers.

1.3. Scope and Limitations

Our results may have been adversely affected by the experiment's limitations, described below.

The research aimed to test ten criteria. Not only is this a large number of variables to test, but the variables themselves were newly created during the research. Thus the results of this study will need to be verified in further studies by other researchers. The gender criteria used in this study were extracted from literature pertaining to females of all ages. The participants in the study were mostly fourteen years old. We are therefore assuming that the criteria applying to females in general apply also to young girls.

The game created as the experimental tool for the study was simple in concept and implementation. It did not have the support of a commercial firm. Thus the quality of the game may have negatively biased some of the participants' attitudes to the criteria tested. Finally, the sample size of participants used in the study was not sufficient to imply statistical significance. This was due to the study being conducted in a short amount of time.

2. GAME DESIGN

We designed a game concept which aimed to embody each of the criteria listed in the introduction. This concept was enhanced through two participatory design sessions. The game, *MyWorld*, is a thee-tier online system. This section initially describes the game features, after user participation had been conducted. It then outlines the design of the system, including considerations found in literature relating to the database, server and client components.

2.1. Game Description

The client's graphical user interface is shown in figure 1 below.

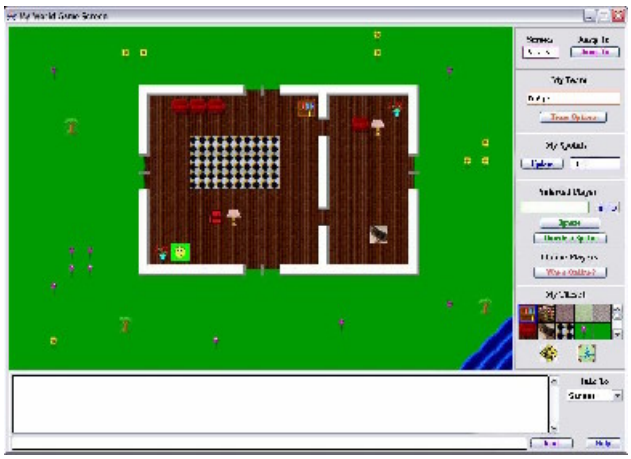

Figure 1. A screenshot of *MyWorld*.

Each of the game features available in the interface are described below. This description is vital to understand how the gender game criteria were embodied in the game. This discusses after the features given below.

2.1.1. Game Features

Players in the game can customise their avatars. They are able to set their details, including name, age, location, school and interests. Furthermore, players can set their portrait to any image they want. Users can also submit tile images to be shown on the landscape in a similar way to portraits, allowing them to create any type of landscape they want. The game world is divided into screens. Each screen is a rectangle of 28 by 19 tiles. Players can move from one tile to another or can jump from one screen to any another. Players can type chat messages that have three different destinations: every player in the player's screen; every player in the player's team anywhere in the world and a selected player anywhere in the world. Additionally, a player can type a ` character before a message to describe an action in a message. All users have a set of tiles on their computers. These include tiles downloaded with the game and any tiles they have created themselves.

Players can place tiles on the landscape if they have sufficient *spoints* (discussed below). Tiles can be placed only on a block owned by the user or on an unowned block. Each player is given sixty *spoints* (social points) when they begin the game. *Spoints* are used to limit the number of tiles and images a player can add to the game. *Spoints* are gained through social actions and the popularity of a player's landscape. *Spoints* are lost when a user places tiles, adds tile images to the game, joins a team (discussed below), creates a team or is ignored by another player. Players can band together in teams. It costs a player *spoints* to join and create teams; but the *spoints* are an investment, as being in a team increases the number of *spoints* a player gains. Each player in a team receives *spoints* when any player in the game walks across a tile belonging to a player in the team (with the owner of the tile thus gaining double the amount of points she would get if not in a team).

2.1.2. Embodiment of Criteria

We designed the game features described above to embody the gender gaming criteria. How each criterion is embodied in the game is listed below:

- No over-sexualised avatars: Avatars are pictures that are added by the player. Avatars can thus be completely customised by players.
- Indirect competition: Players do not have to compete and can compete indirectly by attempting to build the most popular landscape. Players cannot alter another player's tiles. The *spoint* system also involves social competition where interactions have the indirect effect of altering players' *spoints*.
- Collaboration: Players can join teams to build landscapes and explore collaboratively.
- Emotionally involving: Due to the fact that the game involves representations of real human players and chatting is the major activity, the game is emotionally involving and attractive.
- Non-violent: No violence is included in the game.
- No player death or insurmountable penalties: A player cannot make a wrong decision that causes her to be removed from the game or have to "start over", but is penalised for incorrect social conduct only by a reduction of *spoints* (which can be regained).
- No sexual prejudice: Players can choose to completely ignore other players who harass them. The game interface has no sexually biased language.
- Simple graphics: The graphics are 2D square images. The images themselves can be drawn and submitted by users.
- Creativity: Players can take the basic game structure and create any environment they wish. They can express themselves through speech and landscape design.
- Process orientated: There is no way to win the game and the enjoyment lies in creativity and social communication.

2.2. Database Design

The database was housed in the UCT computer science laboratory, together with the server. The database management system (DBMS) used for this system was *Microsoft SQL Server*. It was chosen based on our familiarity with the DBMS, as well as its powerful SQL dialect, and ability to handle large amounts of data [2]. As Internet connections are notoriously slow, the database was required to be efficient in order to minimise the overall time delay. Database optimisation was therefore a necessity.

A 2-dimensional coordinate system provided an effective means for maintaining a virtual world. The world was represented as a grid, with each location containing an x and y coordinate. This reduced the importance of advanced spatial indexing techniques. It was considered a valuable addition to the database to store historical information, in order to analyse gameplay patterns. Relations for historical information were therefore incorporated into the database. Finally, an interface class was programmed in Java in order to allow the server to communicate seemlessly with the database. JDBC was considered to be efficient for this purpose.

2.3. Server Design

The purpose of the server was to provide the gaming logic, link the clients and the database, and facilitate communication between the players. To achieve these goals the server would have to be: efficient, to ensure a quick response to players' requests; and robust, to handle the rigours of multiple users playing concurrently. The server was implemented in Java and used TCP/IP sockets to communicate with the clients. It was designed to be efficient, based on considerations found in past research.

The most significant efficiency problem with online game systems is the *server bottleneck*. This term is used to describe the fact that the server is the slowest component of an online game system. Even though a game server's bandwidth and processing capabilities are typically higher than that of a client, the bandwidth and processing power that a server requires increases quadratically per client online [13]. Pure client-server game architectures are based on the publish-subscribe (model-view-controller) design pattern. To avoid requiring a crippling number of updates, each client should be assigned a subscription channel [4]. This means that clients are grouped into channels and handled separately. Separate servers can be assigned to each channel to achieve load balancing of requests. This was accomplished in *MyWorld* by dividing the world into screens. A screen is a channel in this game. Using screens minimises the number of messages transferred between clients. Another technique suggested by [4] to decrease the processing delay at the server, is to store and retrieve only necessary data to and from the database. Thus a client's name and position is cached in the server's memory for the duration of the session, and not requested from the database during every update period. Similarly, chat messages are sent to clients in a screen via the server, without being stored in the database.

2.4. Client Design

A software engineering (SE) methodology that incorporated end-user evaluations was required for the rapid development of the *MyWorld* interface. This was because a limited time was available for the development of the *MyWorld* interface. By involving users in the design process we hoped to ensure that the *MyWorld* interface was usable by our intended end users. The rapid prototyping model [18] was chosen, as it was developed specifically for projects that had a limited development time and included user participation. Once the SE model was chosen, the *MyWorld* interface's development was separated into three phases: the initial prototype development, improving the interface through user participation and heuristic evaluations [12].

Designing and developing the initial working prototype interface without the aid of user evaluations reduced the required time for interface development and design. Involving children would have required additional time, as children often struggle to verbalize their thoughts. This is especially true when abstract concepts and actions are involved [14] [15]. Only upon completion of the prototype would the intended end-users be involved in participatory design to modify the interface. The participants were to play an integral role in the redesign and modification of the *MyWorld* interface. Participatory design would be performed in multiple sessions until the interface was deemed usable by participants. We would additionally perform heuristic evaluations involving experienced HCI practitioners. This would allow us to assess if the interface could be used by all users, and not solely by our intended end-users.

3. METHODOLOGY

Once the game prototype had been constructed, two forms of testing followed: system testing and gender research testing. The system testing aimed to check the efficiency and correctness of the game. The gender research testing involved verifying the embodiment of the criteria in the game and validating the criteria through an online survey.

3.1. System Testing

System testing was conducted on each component of the system. The server testing and client testing are described below. The server testing encompassed performance testing of the database.

3.1.1. Server Testing

Two efficiency tests were performed on the server. An automated client was built to continuously send messages to the server, wait for a response, and record the total time taken per thousand iterations. The two tests were performed using one to eight clients simultaneously. In the first test, the clients requested the number of *spoints* they owned from the server. This test required a single access to the database and a response message sent to the client. In the second test, the clients sent chat messages to all other clients in their screen. This test did not access the database but required the chat message to be multicast to all clients.

3.1.2. Client Testing

Three methods were used to determine the usability of the *MyWorld* interface: cooperative inquiry [3]; constructive interaction [7]; and heuristic evaluations [12].

Because *MyWorld* requires at least two players to be playing the game for collaboration to be experienced by the players, evaluations had to be conducted on more than one user, simultaneously. The cooperative inquiry method therefore had to be amended to use more than one participant. The method best suited to evaluating more than one user, is the constructive interaction method. This method was changed to allow each user to have her own interface. Combining the constructive interaction method with the principles of the cooperative inquiry method allows for more than one subject to be effectively evaluated in a single evaluation. The combined methodology involved the use of two subjects, two note-takers and an *interactor*. One note-taker would record the quotes of the two subjects and the other would record the activities. The quotes and activities would be combined and analysed. The necessary changes required for the interface could then be made.

Finally, three computer science students performed heuristic evaluations.

3.2. Criteria Testing

The criteria were tested in two phases – to check whether they were embodied in the game and to test females' preferences for the criteria.

3.2.1. Criteria Embodiment Testing

Thirteen girls were interviewed to obtain their opinions on whether the criteria were embodied in the game. These interviews were performed after the participants had evaluated the *MyWorld* interface. These participants therefore had thorough experience with the game and hence their opinions were well founded. All relevant principles found in user-evaluation research were incorporated in the interviews. This ensured that we received reliable, unbiased feedback.

The criteria evaluation method took the form of an informal interview. We asked each group of subjects to rate, out of ten, the extent to which each of the criteria was embodied in the game. For a given criterion, ten implied that *MyWorld* fully embodied the criterion, and one implied that *MyWorld* did not embody the criterion at all. In addition, comments made by the interviewees regarding the criteria were also recorded. Based on these, we could suggest ways that the game could better embody the criteria. The comments also showed whether the participants fully understood the meaning of the criterion being tested. For instance, if a subject was asked why she thought the game was violent and replied: "Because you can box people in with your tiles.", the definition would be explained to her again. The participants were interviewed as a group to create a relaxed environment. This promoted discussion among the participants that lead to insightful and accurate feedback.

3.2.2. Criteria Preference Testing

The survey was divided into three phases. These phases corresponded to the ordering scheme proposed by Alreck and Settle [1]. Users were asked relatively undemanding questions for the first ten questions and simple demographics questions for the final six. The initial questions sought to determine a respondent's knowledge of the game and feelings towards computers and computer games in general. Attitudes towards the embodied criteria were pursued during the middle phase. Alreck and Settle's techniques for avoiding bias were incorporated into the survey.

Each user's knowledge of the game must initially be determined in order to avoid bias from users with a limited knowledge of the game. This was identified by [1] as an important consideration. This was analysed according to qualitative data in the opening questions. After all the results from the surveys were obtained, they were analysed as described in the following paragraphs. Each criterion was initially analysed in terms of the number of girls (between ages twelve and sixteen) whose answers agreed or disagreed with the criterion.

These numbers were calculated in one of two ways:

1. Where a Linear-Numeric scale was used in the survey:
   - Answers of one and two are combined and represent a negative feeling.
   - An answer of three represents a neutral feeling.
   - Answers of four and five are combined and represent a positive feeling.

   The number of responses was determined for each of the *three* categories.

2. Where the survey determined only whether a particular criterion was desirable, a "neutral" category was not provided. The number of responses for each of the *two* categories was determined.

A 90% confidence interval was calculated for each percentage. This provided a range in which the true population percentage (for girls between twelve and sixteen) might lie. For this purpose, the population is assumed to be modelled as a Normal distribution. As the sample size is less than 10% of the population size, the population size did not need to be known. [19]. After considering young girls, the tests were extended to all females up to the age of 23. This assessed whether these criteria are appropriate for females of higher ages. The results of the females were subsequently compared with those of the males to determine whether there is a significant difference between the desires of each gender. To test this, the response proportions of each group are compared using the z-test for Normal distribution.

Finally, the results were analysed to determine whether the game is enjoyable as a whole, and whether, this being the case, it is effective in improving the feelings and attitudes of young girls towards computers. In order to analyse this, a comparison was made between the percentages of users from each category who liked the game, disliked the game or were neutral towards the game. In addition, the average ratings towards the game, provided by each category of respondents, were compared. To test whether these ratings are significant, one-sided t-tests were conducted. These compared the average rating to a rating of three. If the rating was found to be significantly above three (at the 5% significance level) then we could be highly confident that the game would be enjoyed by the population.

4. RESULTS

This section details the results of the study, including both the performance of the three-tier system and the gender gaming criteria experiment.

4.1. Server Results

The results of the two server performance tests outlined in the methodology section are given in this section. The first test, the PointsMessage test, included a database access and returned a message to the client. Figure 2 below shows the number of seconds it took to send and receive one thousand PointsMessages per number of clients online. Each client added to the system caused a delay increase of one second on average. It was unexpected that the delay decreased when a second client was logged on. From the second client to the eighth the delay increased in a monotone fashion. This is shown by the linear trendline appended to the graph. The total time to send and receive one thousand PointsMessages when eight clients are online is 8.3 seconds. This is equivalent to 120 messages per client serviced every second.

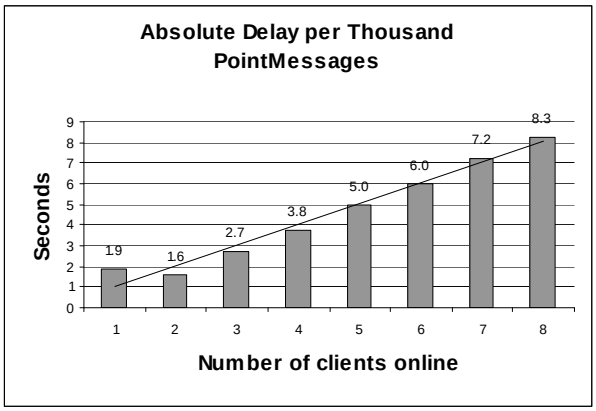

Figure 2. Absolute Delay per Thousand PointMessages

The second test, the ChatMessage test, made no use of the database and forwarded the message received from one client to all the other clients. Figure 3 below shows the number of seconds it took to send and receive one thousand ChatMessages per number of clients online. Each client added to the system caused an increasing delay to the other clients. It was unexpected that the delay decreased when a second client was logged on (similarly to the PointsMessage test). From the second client to the eighth the delay increased in a monotone fashion. This is shown by the quadratic trendline appended to the graph. The total time to send and receive one thousand ChatMessages when eight clients are online is 46.1 seconds. This is equivalent to 22 messages per client serviced every second.

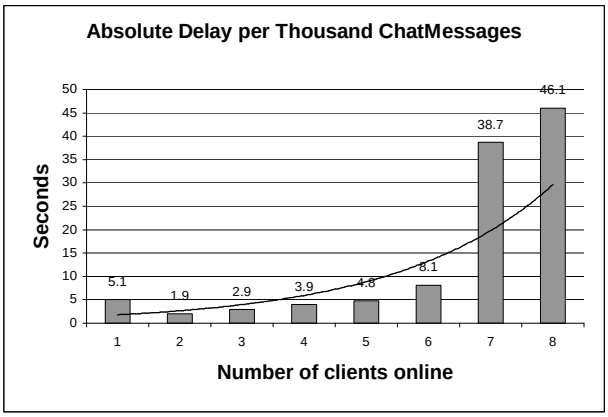

Figure 3. Absolute Delay per Thousand ChatMessages

4.2. Client Results

The user evaluations highlighted various usability problems. These problems were all fixed to the satisfaction of the users. The third user evaluation yielded only two minor usability problems. We therefore concluded that no further user evaluations were required. The heuristic evaluations that followed further verified the usability of the interface by discovering only five minor usability problems. These were also fixed accordingly.

The first user evaluation involved three subjects. As a result of the relatively large size of the group, a range of difficulties were encountered. The *interactor* struggled to assist all three subjects at once. The two note-takers struggled difficult to record the behaviour of the three subjects simultaneously. This resulted in a lack of synchronisation between the quotes and activities obtained by the note-takers. Further difficulties were caused by allowing the subjects to freely explore the game. It was found that the subjects were content just to talk, place tiles and walk around. To ensure that every interface feature was tested we were required to ask the subjects to perform specific tasks.

The difficulties encountered in this evaluation caused us to amend the user evaluation method. Firstly, each evaluator should evaluate only a single subject, recording both their quotes and activities. This would ensure that each subject would receive the full attention of a note-taker and that their quotes and activities would be guaranteed to be synchronised. Secondly, allowing the subjects the freedom to explore an interface was not feasible when the time allowed for our user evaluations was very limited. The solution was to create and use a task-based checklist for future user evaluations. By using a task-based checklist we ensured that every feature of the interface would be tested by the subjects. More informative feedback was obtained in the third user evaluation session than in both previous sessions due to the amended user evaluation method being used.

4.3. Criteria Embodiment Results

This section details whether the participants found the criteria to be properly embodied in the game. The large volume of feedback obtained from the users suggested that performing the interviews with the subjects in a group, was an effective way to obtain their opinions. Being interviewed in a group allowed each subject to speak candidly about their feelings towards the criteria and *MyWorld*. While most criteria were said to be highly embodied in the game (with a rating of above seven out of ten), three criteria were rated below seven. The users felt the game lacked: emotional involvement (rated 6.62 out of 10); collaboration (rated 6.56 out of 10); and creativity (rated 6.8 out of 10).

The lack of emotional involvement was possibly due to the subjects playing the game in a controlled environment. Had there been more time to allow the users to play *MyWorld* without supervision, the users would have been able to form relationships with other users online. This could have led to them experiencing more emotional involvement while playing. The low rating for collaboration can be attributed to *MyWorld* not including team-orientated tasks. There are only two forms of collaboration inherent in *MyWorld*: players belonging to a team gain spoints*; spoints* can be donated to players.

The creativity rating was lower than expected but not excessively low. This was possibly due to: the subjects not being given enough time to create their own landscapes; and the subjects not being provided with a starting area that adequately displayed the creative graphical game possibilities. Providing the subjects with the ability to play *MyWorld* on their own, over an extended period of time, and having the research group spend more time on creating a more creative starting area, would have likely resulted in the subjects experiencing a higher potential for creativity. Aside from the three criteria that received low ratings, the average rating for all of the criteria was 7.97. From this alone we concluded that *MyWorld* embodies the female gaming criteria.

Finally, in every user evaluation, subjects asked if the game had sound. This suggests a possible additional criterion for gender aware games, the inclusion of sound.

4.4. Criteria Preference Results

The table below shows the population confidence intervals for the percentage of females who want a criterion to be embodied in a game.

Table 1: The 90% confident interval for girls (aged between 12 and 16) desiring each criterion

Criterion	**90% confidence interval for the population agreeing with the criterion.**
No over-sexualised avatars	(100%, 100%)
Indirect competition	(40.74% , 92.59%)
Collaboration	(7.40% , 59.26%)

Emotionally involving	(40.74% , 92.59%)
Non-violent	(17.11% , 71.77%)
No player death or insurmountable penalties	(28.23% , 82.89%)
No sexual prejudice	(28.23% , 82.89%)
Simple graphics	(61.04% , 100%)
Creativity	(17.11% , 71.77%)
Process orientated	(100% , 100%)

The intervals above vary largely for most criteria. This is as a result of the small sample size used. The intervals do, however, present "no over-sexualised avatars", "simple graphics" and "process orientated" as desirable with 90% statistical confidence. These intervals have a minimum value above 50%.

The following results are based on the responses of the sample. The criteria found to be liked by the majority of girls within this age groups were: 1, 2, 4, 6, 7, 8 and 10. Criteria 1 and 10 were supported by all girls. Criterion 5 had a high number of users who felt neutral towards it. There were, however, more users in favour of the criterion than opposed to it. Criteria 3 and 9 had more users opposed to them than in favour of them. It was, however, acknowledged that the manner in which these criteria were tested did not necessarily identify users as being opposed to a criterion, but rather that the criterion is not a priority.

5. CONCLUSIONS

5.1. System Implementation Conclusions

The conclusions that follow in this section were drawn from the system testing in section four.

5.1.1. No More Than Twenty Clients Can Be Accommodated By the Game Simultaneously

The server can service 120 unicast messages per second and 21 multicast messages per second when eight clients are online. This is well beyond the number of messages a human player would be able to send per second. It is particularly reassuring that the database accesses required for the PointsMessages are not an unacceptably large overhead. Consider the linear relationship shown by columns 3, 6 and 8 of figure 3. If it is assumed that delay doubles for every two clients added to the system then it can be extrapolated that when twenty clients are online the delay will be 448 seconds. This is equivalent to just under two messages a second. Few human users would want to receive messages at less than this rate. Thus, based on the results at hand, it can be concluded that *MyWorld* can host a maximum of twenty clients at one time. This may be an overestimate, as clients playing over the Internet would not be able to send messages to the server as quickly as ones running on the network. Thus the server might be able to cope with a few more than twenty clients simultaneously. *MyWorld* could be upgraded to service more clients based on this information. Since database accesses cause a linear (and therefore negligible) increase in delay, the game could be altered such that each screen was hosted by a different server (with a maximum of twenty clients per server). This would mean that one hundred different servers would be needed to host the game commercially and that the database would have to service concurrent access from multiple servers.

5.1.2. The Server Can Service Two Clients Faster Than One

Figures 2 and 3 show a surprising anomaly, that the server appears to service two clients faster than one. We are at a loss to explain this, as it cannot be accounted for by any of the system's code. One possibility is that when two clients are interacting with the server it has to do more processing and network communication than when just one client is interacting with the server. This might mean that the operating system running the server gives more operating time to the server than it previously did, when the server had only one client. Thus the server may have more time to access the socket and to process requests and therefore respond faster to two clients than just one.

5.1.3. An Online Multiplayer Game in Which Users Can Add Their Own Images is Feasible

The development and testing of *MyWorld* has shown that creating an online game in which players can create their own images and place them in the game world is feasible. *MyWorld* is such a system that was implemented successfully and easily. The game itself was programmed in one month. Thus a commercial game of higher quality could be created in a slightly longer time. A game such as *MyWorld* does, however, have a need for high capacity networks. Since users can create their own images, not every user will have all the images in their client files at any given time. Downloading these images (even though they are small) takes many times longer than downloading a message to update a client that contains only the name of the image to be shown on the screen.

5.1.4. The New User Evaluation Method Used in This Study was Highly Effective

More informative feedback was obtained in the third user evaluation session than in both previous sessions, due to the amended user evaluation method used. This would suggest that the evaluation method is a highly effective method of obtaining feedback from child subjects using a collaborative system.

5.2. Gender Experiment Conclusions

The conclusions that follow in this section were drawn from the gender research results in section four.

5.2.1. The Criteria Were Embodied in the Game

MyWorld embodied the criteria sufficiently for their validity to be tested by means of the online survey. However, the three criteria with low ratings may have adversely impacted upon the results of the survey.

5.2.2. All But Three of the Criteria Should Be Incorporated in Gender Aware Games

The ten female gaming criteria proposed in this study were shown to be valid. All but three of them should be incorporated in computer games aiming to be gender aware. Our conclusions regarding each criterion, based on the results in section four, are discussed below:

- A low proportion of girls indicated a desire for collaboration and creativity.
- Users are indifferent to violence in games. More girls, however, favoured non-violence over violence.
- Indirect competition and emotional involvement are desirable. Most girls did not desire more competition than is present in the game and most expressed a desire for emotional involvement in computer games.
- Games should not have player death, or sexual prejudice. Most girls do not favour their characters dying, while a majority indicated that they would ignore players who made sexually offensive comments.
- Simple graphics are adequate, as most girls were content with the level of graphics used in the game.
- All the girls were happy with the process orientated nature of the game, and all games that girls indicated they prefer are largely process orientated.
- Games should never have over-sexualised avatars. No girl selected an over-sexualised image as a portrait.

These results confirm the validity of seven of the ten proposed criteria. Collaboration, creativity and non-violence are not necessarily desirable criteria for a game targeted towards girls. The same criteria are relevant to females of a larger age range (between 12 and 23), except creativity is more desirable.

5.2.3. Males Have Different Gaming Preferences To Females

The survey identified the following differences when comparing the results of males and females. Males prefer games with more violence. Males are less opposed to their avatar dying in a game. Males enjoy creativity in a computer game less than females. Males prefer direct competition. Males are less sensitive to sexual prejudice.

5.2.4. The Game Was Enjoyed By the Girls Surveyed, But It Did Not Change Their Attitudes Towards Computers

The game was found to be enjoyable by the girls between ages 12 and 16, as well as females between the ages 12 and 23. The game was not enjoyed by males. *MyWorld* was therefore found to be less gender neutral than it is female orientated. *MyWorld* did not change the opinion of girls towards computers or computer games. All girls had however played computer games before and did not have negative feelings towards computers.

5.3. Future Research

The criteria, which attempt to be a summary of most gender gaming research, require further testing. This includes testing with larger sample sizes and using different games.

An alternative to surveys in future research would be to analyse the behaviour of players within the game, as well as the demographics of players who played the game. This could serve as a better indication of the desirable criteria, than a subjective survey. As a result of limited time this analysis was not performed in this study. If the game becomes widely used on the Internet, the database structures are in place to perform these analyses.

Future academic research should include extending the concept of users being able to dynamically alter an online virtual environment from 2D tiles to 3D objects. 2D images are small in comparison to 3D objects and are quicker to manipulate and send across a network. Additionally, updating players' views dynamically would be far more complex than in *MyWorld.* The database and user interface components would also require more research to support extending the game to three dimensions.

If the game were to be implemented commercially in its current form a few improvements would need to be made. Security would have to be added to the socket communications. Security was unnecessary when the game server was used only as a tool to test the gender gaming criteria, but would be essential if the game were used by hundreds of players expecting a high quality of service. Additionally, professional artists would have to be hired to create aesthetically pleasing tiles and portraits to be bundled with the game. Though many users enjoy creating their own images, they still want to have appealing pre-drawn images available.

All three of the conducted user evaluations seemed to identify that the criterion of sound was essential to their gaming experience. As a result, future research should be performed to verify that sound is an additional criterion that is required by female gamers.

The criterion of "simple graphics" should be further investigated. The *MyWorld* interface could be upgraded from a two-dimensional game with "simple graphics" to a three-dimensional game with "advanced graphics". Females' enjoyment of the three-dimensional game could then be compared to their enjoyment of the two-dimensional game.

Whether the amended child evaluation method described in this article is more effective than current evaluation methods should be investigated. This would require performing further user evaluations, using existing child user evaluation methods and the amended evaluation method, on various different collaborative systems developed for children. The results of these tests could then be used to compare our user evaluation method to existing methods.

6. ACKNOWLEDGMENTS

Our thanks to our supervisor, Sonia Berman. HCI advice was given by Gary Marsden. We also appreciate the invaluable contributions of the participants.

7. REFERENCES

[1]. Alreck, P. L., and Settle, R. B. *The Survey Research Handbook*. Second Edition. USA, Irwin McGraw-Hill, 2002.

[2] Chigrik, A. *The comparison of SQL Server 2000 with MySQL v4.1*. (http://databasejournal.com/features/mssql/article.php/3087841)

[3] Druin, A. Cooperative Inquiry: Developing New Technologies for Children with Children. Human Factors in Computing Systems: CHI 99. ACM Press, 1999.

[4] Fiedler, S., Waller, M., and Weber, M. *A Communication Architecture for Massive Multiplayer Game, NetGame.* Braunschweig, Germany, 1999.

[5]Graner-Ray, S. *Gender inclusive game design: Expanding the market*. Charles River Media, Massachusetts, 2003.

[6]Gürer, D., and Camp, T. *An ACM-W Literature Review on Women in Computing.* SIGCSE Bulletin, Vol. 34, No. 2, 2002.

[7] Kahler, H., Kensing, F., and Muller, M. *Methods & tools: constructive interaction and collaborative work: introducing a method for testing collaborative systems*. ACM Press. 1999

[8] Kiesler, S., Sproull, L., and Eccles, J. *Pool Halls, Chips and War Games: Women in the Culture of Computing*, SIGSCE Bulletin, Vol. 34, No.2, 2002.

[9] Kirk, M., Zander, C. *Bridging the digital divide by co-creating a collaborative computer science classroom. JCSC*, Vol. 18, No. 2, 2002.

[10] Mine, M., Shochet, J., and Hughston, R. *Building a massively multiplayer game for the million: Disney's Toontown Online.* ACM Computers in Entertainment, Vol. 1, No. 1, Article 6, 2003.

[11] Natale, M. *The effect of a male-oriented computer gaming culture on careers in the computer industry*. Computers and Society. 2002.

[12] Nielsen, J, and Molich, R. *Heuristic evaluation of user interfaces*, Proc. ACM CHI'90 Conf. (Seattle, WA, 1-5 April) (pp. 249-256), 1990.

[13] Pellegrino, J., and Dovrolis, C. *Bandwidth requirement and state consistency in three multiplayer game architectures*. NetGames. Redwood City, California. 2003.

[14] Piaget, J. *Psychology and Epistemology: Towards a theory of knowledge*. New York: Viking Press. 1971.

[15] Piaget, J. *To understand is to invent: The future of education*. New York: Grossman. 1973.

[16] Rowell, G., Perhac, D., Hankins, J., Parker, B., Pettey, C., and Iriarte-Gross, J. *Computer related gender Differences.* SIGCSE. 2003.

[17] Seay, A., Jerome, W., Lee, K., and Kraut, R. *Project Massive: A study of online gaming communities*. CHI. 2004.

[18] Tripp, S. D., and Bichelmeyer, B. *Rapid protoyping: An alternative instructional design strategy*. Educational Technology, Research and Development, (pp. 31-44), 1990

[19] Underhill, L., Bradfield, D. (2000) *IntroSTAT*. Second Edition. Cape Town, Juta & Co, LTD.