An Exploratory Study on Nepalese Teenager's Visual Recognition and Preferences in Serious Games

Hua Bai
School of informatics,
University of Skövde
Skövde, Sweden
hua.bai@his.se

Björn Berg Marklund
School of informatics,
University of Skövde
Skövde, Sweden
bjorn.berg.marklund@his.se

Ran Zhang
School of informatics,
University of Skövde
Skövde, Sweden
ran.zhang@his.se

Abstract

In serious game development, effective communication through both languages, sounds, and icons can be crucial for a game to have its intended impact. While this is also true for entertainment games, serious games have added layers of challenges as they are a) often played by audiences outside of the "typical" game ecosystem, and b) miscommunication can lead to players missing important lessons or even learning incorrect information. When a serious game is intended to be used in different parts of the world, however, clear visual communication gets an added layer of complexity: culturally informed symbol interpretation and visual preference. In order to examine how these might affect players' experiences when playing serious games, this paper presents the results of a mixed-method study conducted in two schools in Nepal. The study included 10 participants, between 13-16 years old, who played a prototype of a mobile game currently in development, which has the purpose of teaching young players about food nutrition and healthy habits. After playing the prototype, they took a short survey where they were asked to identify different food types, and they were also interviewed to discuss their opinions of the game's visual style. The results of the study indicate that, while higher fidelity images were much easier to correctly identify by the participants, the participants' preference for visual fidelity varied to a large degree.

Keywords— serious games, educational games, visual fidelity, visual preference, healthy habits, art direction

I. INTRODUCTION

The development of serious games is a complex interdisciplinary practice. A broadly recognized component of this complexity is the difficulty of merging entertaining game mechanics with accurate subject matter representations [1]. Another challenge is understanding the characteristics of the game's target audience. It is well established that different players all have individual levels of gaming literacy, which can be affected by myriad factors (age, upbringing, technology access, etc.). These differences mean that a game will be experienced differently by individual players - and creating designs that accommodate heterogeneous audiences complicates the development of serious games even further. There is plenty of research done outside of the field of game studies and serious games that has examined how cultural contexts shape peoples' preferences and interpretations of different shapes, styles, and colors [2,3,4]. In a cross-cultural study of visual preferences in website design, for example, showed that preferences in colorfulness and visual complexity was impacted by gender, socio-cultural contexts, and levels of education [5].

However, within game research, pros and cons of different visual styles is rarely examined. While there are examples of research relating to visuals in serious games [6], these focus on high fidelity as an across-the-board goal, rather than examining whether individual players' prefer or benefit from it. As games (and, by extension, serious games) are a highly visual medium, understanding these individual differences is crucial when creating games for multi- or cross-cultural play contexts.

For entertainment game developers, creating games that players find visually appealing is, naturally, important since it impacts engagement and marketability. But for serious game developers, just as it is important to balance entertaining game mechanics with educational content, it is also important to balance visual appeal with visual clarity. Lack of clarity introduces risks of conveying subject matter contents in confusing or misleading ways, but an over-emphasis on clarity over novel artistic impressions can make serious games seem drab and thus lead to a decrease in adoption and retention.

In this paper, an empirical study on Nepalese teenager' recognition of different visual styles, as well as their preferences of visual styles, is presented. During the study, participants between ages 13-16 were observed while they played an early prototype of a serious game about food nutrition and healthy lifestyle choices (*Happy Heart*). The observations particularly focused on their interpretation of different food items within the game. The participants' performance in the game, however, wasn't the primary point of interest for the study design. After the play session, the participants went through a quiz where they needed to identify different food items drawn in different styles, and were also interviewed about their thoughts about the game's visuals as well as their video game habits and media preferences. We also wanted to investigate whether the participants had different preferences in visual styles in different media. For this reason, we also asked the participants what type of style they'd prefer to see in games, and which style they preferred to see in textbooks.

It is important to note that this study serves as a pilot for a large-scale four-year project, and should be read as such - the limitations of the study will be discussed at the end of the Results section.

II. RESEARCH METHOD

A. Participants and research context

The empirical data was collected through 10 play-tests and interviews in Nepal between the 21st and 23rd of June. Teachers in both a government school, and a private school, Shree Padma and Golden Sungava respectively (see Fig. 1) hand-picked five students from 13 to 16 years old (Grade 8 to 10). Five participants were female and five were male. All the play testing and interviews were conducted on-site and in

English. As the English proficiency of the participants varied, a Nepalese research assistant was involved as an interpreter during interviews. During the gaming section of the study, the participants sat by a researcher who acted as the game's "computer", responding to the participants' inputs/game choices. For example, the researcher would give the participants points for correct choices. An observer sat across the table and recorded their observations, and the Nepalese interpreter was seated by the observer. For the subsequent interviews, the participant was seated beside the interviewer and interpreter in a classroom.

Fig. 1. Interview and paper-based prototype play testing setup. The setup environment in a government school (top) and private school (bottom)

As for the characteristics of the participants - they all had access to digital media, all of them had their own phones or access to their parents' phones. Six participants have computers, and three have tablets. When asked the average number of hours per week they had played any type of video game the past year, one subject reported more than 10 hours, two reported 5-10 hours, one reported 1-2 hours, five reported no more than 1 hour, and one reported that they hadn't played any games at all recently. These numbers indicated that the entire sample population has some extent of interacting experience on digital devices, but varying levels of gaming interests. The diversity of playing frequency also provides the case study with a more realistic and broad analysis of the proposed aim. In certain interview responses, it became clear that previous media exposure influenced the participants' visual preferences and recognition.

B. Phase 1: Play session

During the play-testing phase, we observed the player's reactions to the different features of the game prototype. The game consisted of two different mini-games (Fig. 2). In the first one, they were tasked with placing different food types in their correct places in a food pyramid. In the second one, they placed food items on a seesaw. The different food items had different "weight" (based on their sugar contents), and the goal was to balance the items correctly on the seesaw. When they finished the two levels of the game, we asked five open-ended questions regarding the prototype.

Fig. 2. The two mini-games of the paper prototype of Happy Heart

It is important to note that the students' performance, or attitudes regarding the game prototype's entertainment value, is not the primary focus of this study. Since this study doesn't aim to evaluate game experiences, but rather investigate visual preferences and recognition, the play experience is secondary and won't be subject to much analysis in the remainder of the paper. We still wanted to describe the play session and game, however, since it was a component of the participants' journey through the overall study.

C. Phase 2: Interview and recognition test

To better understand the participants' preferences of visual styles in games, as well as their ability to interpret different visual styles, semi-structured interviews were held after the play-session. The interview was divided into two different sections. First, we asked questions about media habits (e.g., how often they played games, what games they preferred to play, their favorite tv-shows, their attitudes towards educational games, etc.). Second, we designed three sets of different common food items (apple, onion, and roti - which is a type of flatbread) with five different levels of fidelity to test visual recognition and preference (Fig. 3).

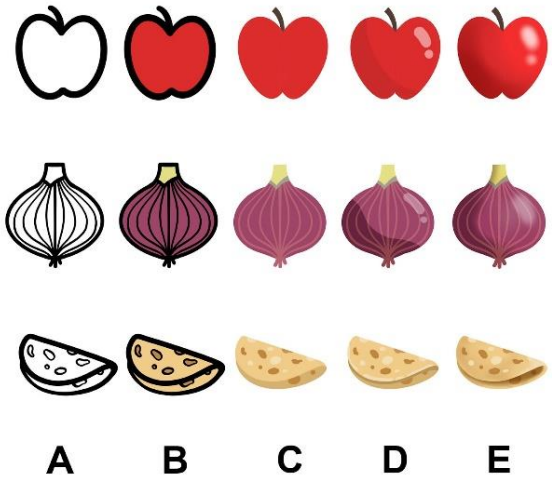

Fig. 3. testing pictures with different levels of fidelity. A: No color, only black outlines of the object. B: black outlines of the object with flat colors. C: no outline, only the shape of the object, with slightly more color. D: no outline, the object has color, sharp shading. E: No outline, the object has color and more realistic soft shading.

Previous research has shown that abstract and symbolic game visuals seem to be preferable for learning and providing an engaging experience [7]. Human visual perception unconsciously ignores the least visible parts when presented with overloaded information. Thus, we framed the graphic style into abstract style and created objects with five different levels of fidelity. The selection of these particular food items is based on visual representation difficulty and cultural relevance. These food items are common in Nepalese cuisine, and the intention was to limit the potential of visual recognition being affected by a lack of previous exposure to the items rather than the visual style.

In order to test visual recognition, we showed the different styles of the food items on a mobile screen, and asked participants if they could identify them. In order to also see what visual style they preferred, we then showed them all versions of the items at the same time so that participants

could show which ones they liked, and discuss why they liked them. (See Fig. 4)

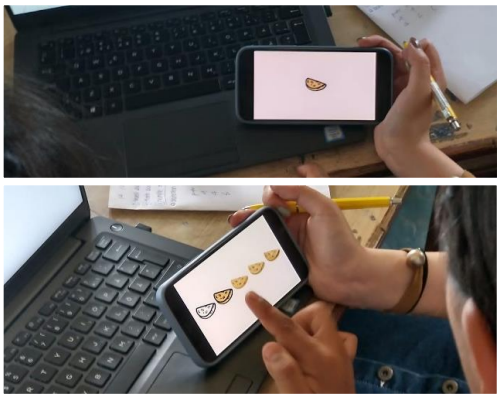

Fig. 4. The participants were shown the five levels of fidelity of the different food objects individually (top), to see if they understood what the picture conveyed. Later, they were shown all five levels of fidelity combined (bottom), to discuss which one they liked the most, and why they found some of them easier/harder to identify.

D. Data collection

In play-testing, the score of the game, players' reactions, and players' answers in the mini-games were recorded on a notepad. Since this was an exploratory study, and conducted within a different cultural context than what we're accustomed to, we wanted to prioritize giving participants ample room to express their experiences and thoughts freely. Due to this, we adopted a semi-structured interview method following the guidelines provided by Edwards & Holland [8]. In order to make the interview questions clearer to the participants, all the questions were not just asked verbally but also shown on a laptop screen for them to read at their leisure. Each interview took approximately 25 minutes, and were recorded and transcribed. Further details of the interviewees are summarized in Table I.

TABLE I. DETAILS OF INTERVIEWEES. THE PARTICIPANT CODE IS BASED ON THEIR GRADE, THE NAME OF THEIR SCHOOL, AND THE NUMBER OF STUDENTS

	Code	***Gender***	***Age***	***Grade***	***School***
1	9SP1	F	15	9	Government
2	8SP2	M	14	8	Government
3	9GS1	M	15	9	Private
4	9GS2	F	14	9	Private
5	8GS3	F	13	8	Private
6	10GS4	M	15	10	Private
7	10GS5	M	16	10	Private
8	8SP3	F	14	8	Government
9	10SP4	M	14	10	Government
10	10SP5	F	14	10	Government

E. Data analysis

First, we transcribed all interview records of the manuscripts. Second, we created codes based on the relevant game and visual concepts and observations regarding game activities. Third, we coded all the manuscripts, sentence by sentence. Then, according to the similarities and differences, we clustered the compiled codes in order to create categories for game visual communication.

III. RESULTS

A. Visual recognition

When the participants went through the recognition test of the items in different visual styles (see Fig. 4). As shown in Table II, all of the participants could immediately recognize the apples and onions at fidelity level A. The recognition of the roti was, however, more varied. Three out of the ten participants couldn't recognize the roti at all, across all visual styles. Among the participants that could ultimately identify the roti, some of them still hesitated and had to think about their answers for a relatively long time.

TABLE II. NUMBER OF PARTICIPANTS THAT COULD RECOGNIZE ITEMS AT THE DIFFERENT VISUAL STYLE LEVELS

Visual style	***Apple***	***Onion***	***Roti***
A	10	10	1
B	0	0	1
C	0	0	1
D	0	0	3
E	0	0	1
Couldn't recognize	0	0	3

After the participants had gone through the recognition test, they were asked why they felt the different levels were easy or difficult to identify. In response to the question of which one do you think is most easy to recognize, comparing the data of three sets of different objects, participants tended to recognize images with the highest degree of authenticity when the "difficulty" of the object increases.

Without the color, I could not really understand what this was. But after it was in color I started understanding that it was like roti. (8GS3)

(Roti A) There is not any color. So, I think it looked like cheese. (9GS2)

The second one (B) has colors. That made it feel more like a roti. The third one (C) is clearer. In the fourth one (D), the shadows added more realism. The fifth one (E) was easier to recognize. (10GS4)

(Roti E) This part of the roti has a little shadow, so we can know that this is a little bit like a folded roti. (10SP4)

As these examples above show, complex shapes like roti can be mistaken for other objects when shown without color, especially when they are similar in shapes. Table III shows all different answers the participants had to the different style levels of the roti. As the fidelity of the image increased, more and more participants recognized the image as it was meant to be. One particularly large jump in recognition happened in style D, where shading was first introduced in a more realistic way.

TABLE III. RECOGNITION OF FIVE DIFFERENT FIDELITY OF ROTI

	Style A	*Style B*	*Style C*	*Style D*	*Style E*
1	Cookies	Cookies	Cookies	Cookies	Folded cookies
2	Pumpkin	Pumpkin	Cookies	Roti	Roti
3	Watermelon	Don't know	Don't know	Roti	Roti
4	Don't know	Cheese	Cheese	Cheese	Roti
5	Taco	Don't know, Taco	Roti	Roti	Roti
6	Watermelon	Roti	Roti	Roti	Roti
7	Pizza	Pizza	Cookies	Cookies	Pizza
8	Roti	Roti	Roti	Roti	Roti
9	Don't know	Cheese	Cheese	Roti	Roti
10	Cookies	Cookies	Dried Snacks	Dried Snacks	Dried Snacks

B. *Visual preferences for recognition*

Even though participants had an easy time recognizing the more "natural" food items (apple and onion), we wanted to see what visual style they thought was most recognizable. The results can be seen in table IV. Four participants chose Apple E, three participants chose Apple D, two chose Apple C, one chose Apple B, and no one chose Apple A. The recognition of simple objects is easy even when the fidelity is low. For example:

(Apple B) Because it looks easy to understand that this is an apple. This is a picture that nursery teachers show us when we were nursery. (8GS3)

(Apple D) It looks quite similar to the real apple (10SP5)

Some of them reported that a simple outline helped them to recognize the items:

(Onion B) the outline is also clear and it is easy to understand for kids as well (10GS4)

TABLE IV. DISTRIBUTION OF PARTICIPANTS' PREFERENCES OF DIFFERENT VISUAL STYLES (WHAT THEY THOUGHT WAS EASIEST TO RECOGNIZE)

	Apple	**Onion**	**Roti**
A(%)	0	0	0
B(%)	10	20	20
C(%)	20	10	0
D(%)	30	20	0
E(%)	40	50	80

No one chose Apple A, Onion A, or Roti A. Although they could recognize the object without color, most did not state it as the easiest, and that they would further benefit from a more complex style.

These responses echoed observations we made during the play session. All food items in the paper-based prototype were similar to visual style D (in color, without an outline, and sharp shading). In the game, participants could easily recognize almost all "natural" food items. It was, however, difficult for them to recognize many cooked food items. For example, a pasta item was seen as chow mein or thukpa, and pudding was recognized as a cake. All of the participants could recognize a can of coke and purple grapes. However, they mentioned that a bottle of coke and green grapes is the most common one.

C. *Differences in preferences between textbooks and games*

When asking participants about their preferred visual styles in games and textbooks, the answers tended to be relatively similar - where similar styles were preferred both in books and in games (see Table V). They weren't, however, identical. There was a big difference in how the participants wanted an onion to look, and the participants were evenly split when it came to the roti. Participants were slightly more likely to choose images with low visual fidelity in textbooks when compared to games. Some of the students reported that they had seen similar visual styles in textbooks, thus finding that those styles would be more suitable.

(Apple D) Because I have seen it many times [in a textbook]. (9GS2);

Because (Onion B) fits with the texture of the textbook. (9GS1);

The second one (Roti B), it will be matching the things on our text. (10GS4)

Regarding participants' preference of what visual style they would prefer to see in a game, four participants chose Apple D, four chose Onion E and six chose Roti E. Students tend to choose high visual fidelity in games. Many students mentioned that image E was more realistic and three dimensional.

(Onion D) It looks like a real onion. Something that we prefer is cartoons. (9GS2);

(Onion E) Because it looks like a 3D vision onion. (10SP5);

(Onion E) Because I think the game players would like to have a real looking onion in a game than the fake one. (8GS3);

(Roti E) It's so easy you can take that and these (Roti A,B,C,D) are quite tough I think (8SP2)

TABLE V. PERCENTAGE OF VISUAL PREFERENCE IN DIFFERENT MEDIA

	Apple in book	**Apple in game**	**Onion in book**	**Onion in game**	**Roti in book**	**Roti in game**
A(%)	10	0	0	0	0	0
B(%)	10	20	40	10	50	30
C(%)	20	10	10	20	0	10
D(%)	40	50	20	30	0	0
E(%)	20	20	30	40	50	60

D. Limitation

This study has several shortcomings which makes generalizations ill advised. First and foremost, the sample size is limited, which naturally impacts external validity. In order to fully understand whether the participants' preferences and recognition was culturally informed, more care needs to be done to have a representational sample - where we can be certain that players from different socio-cultural contexts are included. As it stands, the current study can't draw any confident correlations between preference and recognition between different players.

Another limitation is that the objects are drawn according to our own artistic choices. The roti - which differed from the other food objects the most in terms of recognition - might have been drawn in a different position than the participants usually see it. While an apple doesn't leave much room for artistic manipulation, A roti can be cooked in different ways, folded, flat, inflated, filled with other foods, etc. Recognition might very well be tied to how similar it is to how the participants are accustomed to eating roti, rather than visual fidelity in the same way as the onion or apple. It should be noted that none of the participants mentioned this being a reason for them not recognizing it, but we want to acknowledge that it might still be the case.

IV. Discussion

According to the results of the recognition quiz shown in table 2, style E was the clearest image to recognize, followed by style D. In addition, in participants' perspective of easiest recognition, As shown in Table IV, visual style E was the most recognized across all three objects. While the recognition of the apple and onion also increased as the visual fidelity increased, they were more varied. In the participants' own words, the introduction of shadows and light sources were a determining factor of how recognizable an object was. Shortly summarized - and not surprising - when the object looks more realistic, they are easier to confidently identify. The degree in which realism matters, however, seems to vary.

Regarding visual preference, our interview results show that familiarity with the visual object affects users' preferences. Among all the preference results, 2-4 participants in each object mentioned that the reason they liked a picture is based on previous exposure. For example, they have seen the same style in a textbook, or in a game, and thus felt a particular style would look suitable in either medium.

Participants were more likely to prefer style B in a textbook, especially for the onion, and the preferences of the roti was split evenly between B and E. Styles D and E were, overall, a more popular choice in games across all food items.

The data in Table III and Table IV clearly indicate that learners disagree and dislike the picture of a simple outline without colors (style A). In addition, Table III shows the difficulty in recognition of roti. Only one participant correctly recognized style A. The findings of this study fall well in line with a literature review by Bramão that color information contributes to object recognition of drawing [9]. Color knowledge as a result of perceptual experience influences the recognition of objects. These imply that the visual item without color has a negative effect on visual communication.

Based on the data we collected from the recognition quiz, roti as a local food was the most difficult object to identify. Although roti is one of the most common foods in Nepalese daily life, it might not be familiar in their daily media such as books or games. The cognitive performance of localized images is less effective. Visual experience seems to be more significant in recognition than lived experience in this case. Localizing serious game images should take into account players' pre-existing visual experiences. Alternatively, problems with recognition may negatively impact learning outcomes. Does the image need to be more inclined towards basic-level objects in light of this? For example, although a certain type of beer bottle is very common in Nepal, an image of a common beer glass filled with beer might be recognized more quickly. In addition, some objects such as pasta, pudding, and brown rice were difficult to identify in the play session due to the live experience. Thus, the design should balance globalization and localization.

V. Conclusions & Future Work

In this preliminary paper, we have presented an exploratory study on the impacts of visual fidelity in users' experience. Our study shows that recognition performance is best when the object has higher fidelity images. Nonetheless, the data regarding participants' preferences for visual fidelity varied widely. It does not merely refer to the fidelity of the picture but is also influenced by many different factors, such as socio-cultural contexts, and visual perceptual experience. Our research provides preliminary indicators of how different visual representation styles are perceived by players.

In cross-cultural game development, developers should carefully consider visual representation, as it may affect target group adaptation. A better understanding of visual recognition and preferences might reduce risks of visual miscommunication. In serious games, if represented objects that are important for a certain subject matter are misunderstood, learning outcomes might at best be diminished, or at worst the players draw incorrect conclusions from the game. From a more practical perspective, a better understanding of visual communication can help developers make more efficient choices when allocating resource investments between graphic fidelity and other game components. In our future research, we plan to further test different visual representation styles, and introduce a comparative sample with Swedish participants.

References

[1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*

[2] Bahareh Motamed, Richard Tucker, Rose tinted spectacles: Culturally informed differences between Iran and Australia in architect's colour cognition, preference and use, City, Culture and Society, Volume 18, 2019, 100288, ISSN 1877-9166, https://doi.org/10.1016/j.ccs.2019.05.004.

[3] Taylor, Chloe & Grandison (formerly Clifford), Alexandra & Franklin, Anna. (2012). Color Preferences Are Not Universal. Journal of experimental psychology. General. 10.1037/a0030273.

[4] Yokosawa, K., Schloss, K.B., Asano, M. and Palmer, S.E. (2016), Ecological Effects in Cross-Cultural Differences Between U.S. and Japanese Color Preferences. Cogn Sci, 40: 1590-1616. https://doi.org/10.1111/cogs.12291

[5] Katharina Reinecke and Krzysztof Z. Gajos. 2014. Quantifying visual preferences around the world. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14). Association for Computing Machinery, New York, NY, USA, 11–20. https://doi.org/10.1145/2556288.2557052

[6] Chalmers, A. and Debattista, K., 2009, March. Level of realism for serious games. In 2009 Conference in Games and Virtual Worlds for Serious Applications (pp. 225-232). IEEE.

[7] Wouters, P. and Van Oostendorp, H. (2013) A meta-analytic review of the role of instructional support in game-based learning. Computers & Education, 60(1), pp.412-425.

[8] R. Edwards and J. Holland, What Is Qualitative Interviewing?, Bloomsbury Academic, London, UK, 2013.

[9] Bramão, I., Reis, A., Petersson, K.M. and Faísca, L., 2011. The role of color information on object recognition: A review and meta-analysis. Acta psychologica, 138(1), pp.244-253.