
Analytic Evaluation of DEG-7-11

Kornchulee Khanana

University of Leicester
Dept. of Computer Science
LE1 7RH Leicester
United Kingdom
kk207@le.ac.uk

Effie Lai-Chong Law

University of Leicester
Dept. of Computer Science
LE1 7RH Leicester
United Kingdom
lcl9@le.ac.uk

Abstract

We have developed a set of eight design heuristics known as *DEG-7-11* to bridge the gap of existing playability heuristics, which do not take players' age-specific characteristics into account. To validate *DEG-7-11*, Heuristic Evaluation has been conducted by inspecting 15 selected educational games on different topics with the eight heuristics. A statistically significant correlation was found between the players preference ratings of the games and the percentages of the heuristics followed by the games. This result could validate the effectiveness of *DEG-7-11*. An intriguing observation is that none of the games inspected has followed the heuristic about gender-based adaptation. Future research on identifying underlying causes and remedies for improving the situation is called forth.

Author Keywords

Digital educational games, Children, Heuristic evaluation, Player preference, Gender-based adaptation

Introduction

Driven by a goal to develop digital educational games (DEGs) for children aged 7 to 11 to learn basic knowledge of nutrition, thereby increasing their

awareness of healthy eating, we have reviewed existing playability heuristics (PHs) with the aim to identify relevant design guidelines. However, existing PHs (e.g., Desurvire & Wiberg, 2009; Pinelle, Wong & Stach, 2008; Sweester & Wyeth, 2005) are primarily used for entertainment games rather than DEGs and do not take age-specific characteristics of players into account. Based on our understanding of the work of Jean Piaget (Piaget, 1973; see also the review in Hourcade, 2008) and other scholars (e.g., Hesse & Cicchetti, 1982; Iskander, Kapila & Karim, 2010), we are convinced that the cognitive and emotional development of children aged 7-11 entail specific design strategies and approaches. Consequently, we have developed a set of eight heuristics known as *DEG-7-11* (Digital Educational Games for children aged 7 to 11 years) to address the gaps thus identified. The development of *DEG-7-11* has been grounded empirically in our Pilot Study (Khanana & Law, 2013) and in our systematic literature review in psychology, pedagogy and design. Table 1 presents individual heuristics, which are aimed to be intuitive, thereby enabling DEG designers to follow.

To validate the effectiveness of *DEG-7-11*, we have conducted both empirical and analytic evaluations. Specifically, we have developed two versions of a DEG on food groups, which follow fully and partially of the eight heuristics, respectively. The two game prototypes have been evaluated with primary school children aged 7-11. Data of these empirical studies are being analysed. In this paper, we report the analytic

Name	Love It	Rating	Votes
2048	84	9.1	17
Axon	81	8.8	157
Amusix Flute	72	8.2	1077
7 Moves	69	8	772
Perfect Balance 3	64	8	245
Color Traffic 2	64	7.5	1,034
SoloWords	60	7.5	477
Maths Workout 2	58	7.1	995
Double Digits	51	6.5	7851
Word Mountain	51	6.8	784
USA 50 States	50	6.3	292
Race Across the Steppe of Mongolia	40	4.9	132
Are You My Blood Type?	33	4.7	476
Space Words Defense	33	4.6	337
Eyeballing	27	4.1	248

Table 2: Player-based evaluation of each game: Player preference (Love It %), and Overall Quality scores (Rating: from 1 to 10) and Number of players voted (Votes)

evaluation study where we have applied the eight heuristics to a selection of 15 DEGs.

DEG-7-11#1: Instead of setting a goal on performance, such as scoring, like entertainment games, for a DEG, a specific learning goal should be set. Also, the goal should be clear, concise, simple, and presented early in the DEG.
DEG-7-11#2: Children aged 7-11 should be supported by instruction when playing a DEG, the instruction should explain how to get correct answers, but without relying on text-based manual only.
DEG-7-11#3: One single DEG could suit different genders so that children can select or create their own favourite avatar.
DEG-7-11#4: A DEG for children aged 7-11 years should be relaxing to play by having minimalistic interfaces, appropriate speed, and no time pressure.
DEG-7-11#5: A DEG should be separated into multi-levels with initial levels being disguised tutorials, enabling children to practice new information by performing similar tasks.
DEG-7-11#6: A game should incorporate reminders or hints that children can use for recalling information from their memory.
DEG-7-11#7: Animations can influence learning for children aged 7-11 years, especially cartoon-like animations can enhance imagination and fun, resulting in playful learning.
DEG-7-11#8: Rewards and punishments should be provided in the form of in-game feedback interfaces; they are incentives and can inform children aged 7-11 years about their progress and learning.

Table 1. Eight heuristics of DEG-7-11

Methods

Selection of Games

Nowadays, game reviews and game ratings are generally used for judging the popularity of games. We have found a website named Learn4Good¹, which contains free web-based games for children in different school levels, ranging from kindergarten, elementary, middle and high schools. It hosts various interactive learning games on different topics, including mathematics, science, physics, engineering, puzzles, hard brain-teasers, and others. In addition, players can rate a game after playing it with a scale (Figure 1):



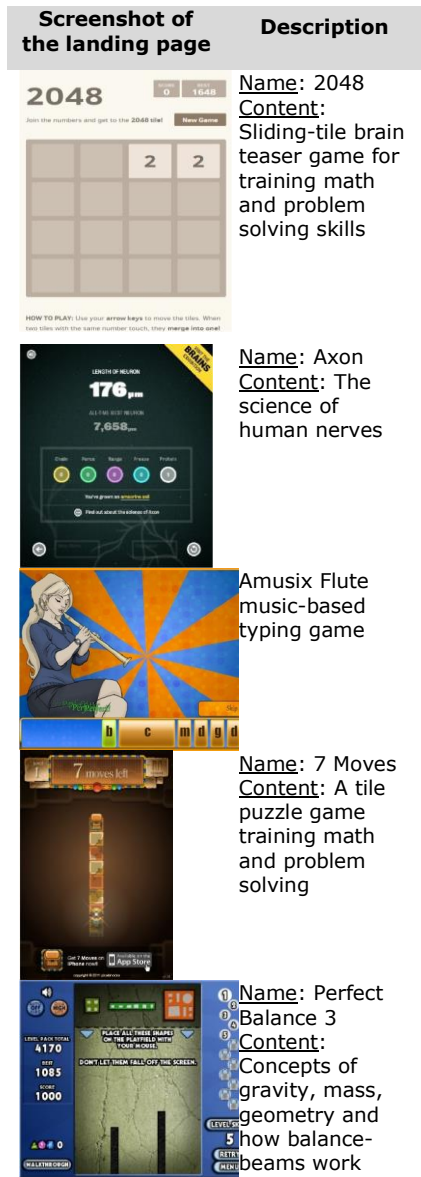
Figure 1: The rating scale used in the website Learn4Good (see footnote 1)

We aimed to validate DEG-7-11 by estimating the extent to which the designs of existing DEGs for children aged 7 -11 are consistent with the DEG-7-11 heuristics. Our research assumption is as follows:

The higher the number of the DEG-7-11 heuristics followed by a DEG, the better its quality and the higher its player-preference rating are.

To verify this assumption, we have selected the games with a range of "Love It %" (Table 2), but limited to a manageable size of 15 games and to those which have been played and voted by a reasonable number of players. To minimize biases, all the selected games are from the category "Educational Games" and run on the same platform (in this case it is Windows OS). No

¹ www.learn4good.com/games



specific contents have been targeted. Details of these DEGs are displayed in Table 2 and Figure 2a, 2b and 2c. In Table 2, the values in the 'Rating' column (with a 10-point quality scale, 1 worst; 10 best) represent the mean averaged over the number of people who have played and rated that game (cf. the "Votes" column). The percentage of 'Love It %' of a game would then be correlated with the percentage of DEG-7-11 heuristics followed by that game.



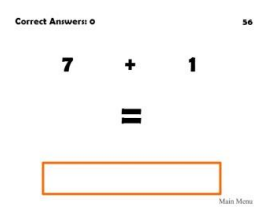
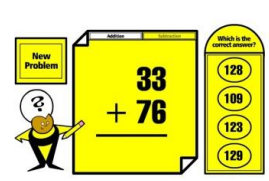

Heuristic Evaluation

A typical Heuristic Evaluation proposed by Nielsen (1994) has been conducted. Accordingly, we have gone through the games interface with two rounds. First, we browsed a game in the first round to know its main goal and its mechanics and to see which features it has. Then we played the game in much more detail, checking each game feature to see if it follows or violates any of the DEG-7-11 heuristics. If a particular heuristic is followed by the game, 'Y' (Yes) is assigned; otherwise, 'N' (No) is assigned. For example, the game named "Axon" is a fast-paced gameplay. The game requires children to have fast reflexes, and accurate mouse-clicking on a Protein Sphere which decreases in size in every second. Its feature violates DEG-7-11#4 (speed and time limit); it is *not* relaxing to play. In contrast with the game named "2048" which has no speed or time limit, children can use their unlimited-time and effort to slide the tiles to combine pairs of identical number tiles to create a tile of the number 2048. As this feature does not violate DEG-7-11#4, then 'Y' is assigned to "2048" and 'N' is assigned to "Axon". The same procedure has been applied to inspect other features of the 15 selected games by using each of the eight DEG-7-11 heuristics.

Figure 2a: Five of 15 selected games (1st to 5th)

Game Name	DEG-7-11 heuristics #								Followed (%)
	1	2	3	4	5	6	7	8	
2048	Y	Y	N	Y	Y	Y	Y	Y	87.5
Axon	Y	Y	N	N	N	N	Y	Y	50
Amusix Flute	Y	Y	N	N	Y	Y	Y	Y	75
7 Moves	Y	Y	N	Y	Y	N	Y	Y	75
Perfect Balance 3	Y	Y	N	Y	Y	Y	Y	Y	87.5
Color Traffic 2	N	Y	N	N	Y	Y	Y	Y	62.5
SoloWords	Y	Y	N	N	N	Y	N	Y	50
Maths Workout 2	Y	Y	N	N	N	N	N	Y	37.5
Double Digits	Y	N	N	Y	N	N	N	Y	37.5
Word Mountain	Y	N	N	N	Y	N	Y	Y	50
USA 50 States	Y	Y	N	Y	N	Y	N	Y	62.5
Race Across the Steppe of Mongolia	Y	Y	N	Y	N	N	Y	N	50
Are You My Blood Type?	Y	Y	N	N	N	N	N	Y	37.5
Space Words Defense	N	N	N	N	Y	N	N	Y	25
Eyeballing	N	N	N	Y	N	N	N	Y	25
No. of games followed that heuristic	12	11	0	7	7	6	8	14	
%	80	73	0	47	47	40	53	93	

Table 3: Percentage of the DEG-7-11 heuristics followed by each of the 15 selected games, as indicated in the column "Followed (%)". The column of DEG-7-11#3 is highlighted to show the unique finding.

Screenshot of the landing page	Description
	<p>Name: Color Traffic 2 Content: Training for being a good and trustworthy traffic control manager</p>
	<p>Name: SoloWords Content: an anagram-type of word game</p>
	<p>Name: Maths Workout 2 Content: Solving mathematical problems under time pressure</p>
	<p>Name: Double Digits Content: Math subtraction and addition</p>
	<p>Name: Word Mountain Content: a fast-paced and innovative typing game</p>

Results and Discussions

The results of Heuristic Evaluation are summarized in Table 3. On average, 54% (SD = 20.4, Range: 25% - 87.5%) of the DEG-7-11 heuristics have been followed by the 15 selected games.

The research assumption stated above implies a significant positive correlation between player preference ratings (measured by the variable 'Love It %') and percentage of the DEG-7-11 heuristics followed by a game (measured by the variable "Followed %". To verify this assumption, we have computed the Pearson correlation coefficient between the two variables. The parametric test has been used, given that the datasets are normally distributed as indicated by the output of Shapiro-Wilk tests.

The result shows that there is a highly significant correlation between Player Preference ratings and percentages of DEG-7-11 heuristics followed ($N = 15$, $r = .753$, $p < .01$). It implies that the higher the number of DEG-7-11 heuristics is followed when designing a DEG, it is more likely that the DEG will have a higher player preference rating.

When computing for each heuristic, 80%, 73%, and 93% of 15 games followed DEG-7-11#1, #2, and #8 respectively. Most DEGs provide "a learning goal", "multimodal instruction", and "reward and punishment feedback" for their players. A learning goal is the source of motivation that inspires children to use their effort to play the game. The multimodal instruction enables children to explore the games themselves. Also, reward and punishment feedback is one of the critical game elements that sustain the player's motivation to play the game. Especially, we have categorized DEG-7-11#1 and DEG-7-11#2 (Table 1) as

essential heuristics in the sense that they must be followed so that a game can be qualified as a DEG. The result of our analytic evaluation can demonstrate that these two heuristics are needed for a DEG design; without applying them a game will probably have lower user preference rating.

In addition, it is found that approximately 47%, 47%, 40%, and 53% of the games followed the DEG-7-11 heuristics #4, #5, #6, and #7, respectively. It can be inferred that the game designers might take into account the psychological frameworks, including emotion (Rolls, 2005; Russoniello, O'Brien & Parks 2009), memory (Baddeley, 1999), cued recall (Moult, 2011) and design guidelines such as animation (Scaife & Rogers, 2005) when creating their DEGs. The games following the heuristics #4, #5, #6 and #7 are multi-levels with initial levels serving as disguised tutorials. The games also incorporate in-game hints to support children's gameplay. The games aim to enhance fun by including cartoon-like animations. Moreover, the games are designed to be relaxing to play. These results tend to support the assumption that if a DEG is designed by following the recommendation heuristics, the game can get higher player preference.

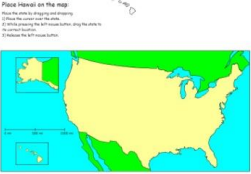
Interestingly, none of the selected 15 games followed DEG-7-11#3 (*One single DEG could suit different genders so that children can select or create their own favourite avatar*). According to Boyle and Connolly (2009), although some guidelines in developing DEGs specify that new learning materials should aim to be gender neutral, traditional computer games are still developed to be more appealing to male than female. A plausible explanation is that developing DEGs that are appealing to both genders requires additional level

Figure 2b: Five of 15 selected games (6th to 10th)

Screenshot of the landing page

Description

Name: USA 50 States
Content: a picture puzzle to learn geography



Name: Race of Mongolia
Content: a quiz game on the history of Mongolia



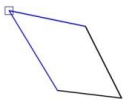
Name: Are You My Blood Type?
Content: to learn which blood types match and which don't.



Name: Space Words Defense
Content: to save the galaxy while practicing typing skills at the same time



Name: Eyeballing
Content: a geometry math game



of complexity and thus additional resources in terms of time and effort. Ideally, DEGs should be adaptive or personalized based on player gender, which is known to be critical factors influencing the motivation to play (Sun & Law, 2010)

Here we discuss some issues pertaining to *gender-based adaptation*. The awareness of gender specificities and gender sensitivity in the game sector is increasing. It is an important step in creating digital games to meet players' special needs. Creating DEGs that are enjoyable and desirable for players of different ages, genders and other characteristics is challenging. A novel approach to learning should benefit learners of both genders (Boyle & Connolly, 2009). Hence, gender-based adaptation framework on a DEG development has been proposed by (Chang, Kuo, Kinshuk, Chen, & Hirose, 2009). The framework especially suits and is needed in the context of DEGs, where a DEG should support learning for all students with equal opportunities. Otherwise, players may mistakenly perceive that the game is not created for them and reject it. Accordingly, different game features, components, and characteristics can be chosen for female or male players. These adaptation variables are derived from the literature. One of main variables is *Game Characters* and *Avatar Preferences*; players should have the chance to select or create their own favorite avatar. It can not only enhance students' motivation and learning performance, but also can improve cost-effectiveness of game development – one game can be adapted for both genders instead of different games developed for male and female players.

For the future work, the reasons underlying the non-compliance with this gender-based heuristic and

appropriate remedies for improving the situation should further be investigated.

Overall, this analytic evaluation study has validated the assumption that DEGs, when following the DEG-7-11 heuristics, are likely to be accepted and enjoyed by their users. However, based only on the player preference ratings, it is not possible to validate if the 15 games evaluated are educationally effective. Results of our empirical user-based evaluation with the two game prototypes will shed light on this issue.

Conclusion

With the increasing use of unconventional educational interventions (cf. chalk-and-talk in classroom) for young children, DEGs are promising tools being developed to promote learning outcome and enjoyment simultaneously. Game designers or UX professionals in the field of Human Computer Interaction need to have guidelines for developing games to ensure quality user experience as well as better learning outcomes. This Heuristic Evaluation study reported in this paper has lent support to the assumption that if game designers disregard some important features, the games might have some deficiencies as identified in the above analysis, leading to its low player preference ratings. In addition, the results infer that DEG-7-11 can be effective guidelines for designers to create a successful DEG.

Figure 2c: Five of the 15 selected games (11th – 15th)

References

1. Baddeley, A. D. (1997). *Human Memory: Theory and Practice*. Psychology Press.
2. Boyle, E., & Connolly, T. (2009). Games for Learning: Does Gender Make a Difference? In T. Connolly, M. Stanfield, & E. Boyle (Eds.), *Games-Based Learning Advancements for Multi-Sensory Human Computer Interfaces: Techniques and Effective Practices* (pp. 288-303). Hershey, PA, USA: IGI Global.
3. Chang, M., Kuo, R., Kinshuk, Chen, G. D., & Hirose, M. (2009). Learning by Playing. Game-based Education System Design and Development: 4th International Conference on E-learning, Edutainment 2009, Banff, Canada, August 9-11, 2009, Proceedings: Springer.
4. Desurvire, H., & Wiberg, C. (2009). Game Usability Heuristics (PLAY) for Evaluating and Designing Better Games: The Next Iteration. In A. A. Ozok & P. Zaphiris (Eds.), *Online Communities and Social Computing, Vol. 5621* (pp. 557-566): Springer Berlin Heidelberg.
5. Hesse, P., & Cicchetti, D. (1982). Perspectives on an integrated theory of emotional development. *New Directions for Child and Adolescent Development*, 16
6. Hourcade, J. P. (2008). Interaction design and children. *Foundations and Trends in Human-Computer Interaction*, 1(4), 277-392.
7. Iskander, M., Kapila, V., & Karim, M. A. (2010). *Technological Developments in Education and Automation*. Springer.
8. Jacob, S. H. (1984). *Foundations for Piagetian education*. University Press of America.
9. Khanana, K., & Law, E. L. C. (2013). Designing children's digital games on nutrition with playability heuristics. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 1071-1076). ACM.
10. Moul, M. (2011). Cued Recall. In J. Kreutzer, J. DeLuca & B. Caplan (Eds.), *Encyclopaedia of Clinical Neuropsychology* (pp. 751-752). Springer New York.
11. Nielsen, J. (1994). Usability inspection methods. In *Conference companion on Human factors in computing systems* (pp. 413-414). ACM.
12. Piaget, J. (1973). *The Child and Reality*. New York: Grossman, 1973.
13. Pinelle, D., Wong, N., & Stach, T. (2008). Heuristic evaluation for games: usability principles for video game design. In *Proc. Conference of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, Florence, Italy.
14. Rolls, E. T. (2005). *Emotion explained (Vol. 22)*. Oxford University Press Oxford.
15. Russoniello, C. V., O'Brien, K., & Parks, J. M. (2009). The effectiveness of casual video games in improving mood and decreasing stress. *Journal of Cyber Therapy and Rehabilitation*, 2(1), 53-66.
16. Scaife, M., & Rogers, Y. (2005). External cognition, innovative technologies and effective learning. *Cognition, Education and Communication Technology*, 181-202.
17. Sun, X., & Law, E. L. C. (2010). Towards a structural model for intention to play a digital educational game. In *Transactions on edutainment IV* (pp. 44-55). Springer Berlin Heidelberg.
18. Sweetser, P., & Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *Comput. Entertain.*, 3(3), 3-3.