
Get the Picture? Evaluating Interfaces through Children's Drawings

Cristina Sylla

University of Minho
EngageLab
Guimarães, Portugal
sylla@engagelab.org

Abstract

We conducted a study to determine whether it was possible to evaluate the usability of a children's interface just by looking at their drawings, uncovering indicators that would reveal the degree of success of the interaction.

Two groups of children aged between four and five years old were exposed to two versions of a computer game. In the regular version the game worked as expected, in the other version the mouse would stop functioning during random periods of the game play. The drawings made by the children after the game were analyzed by three evaluators to determine if they corresponded to the interaction with the regular or the broken game. The results show that in this specific study the decoding of children's drawings made after their interaction was clearly insufficient to assess the usability of the interface, and that further research is needed.

Keywords

Children, Technology, Drawings, Evaluation, Usability.

ACM Classification Keywords

H5.2. User interfaces: Evaluation / Methodology.

General Terms

Human factors

Copyright is held by the author/owner(s).
CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.
ACM 978-1-60558-930-5/10/04.

Introduction

Children, from an early age, are ever more exposed to technology, either for leisure or education. Nevertheless the evaluation of technology with children aged four to five years old poses some difficulties since they might have difficulties to express themselves clearly through words; in addition the great majority can neither read nor write. Given the previous point, most evaluation methods are better suitable to be used with older children. Younger children find it often easier to express their feelings through drawings rather than through words. Drawing is one of the main activities carried at pre-school and it is often used to assess children's involvement after certain activities [3]. The work presented here investigates the use of drawings as a method to convey information on the usability of a children's computer game.

Previous Work

Drawings have an historical tradition as a method to evaluate cognitive development. Lowenfeld's concepts, especially is visual-haptic theory, still continue to influence a number of research areas such as: art education, child studies and psychology. Lowenfeld sees drawing as a process that children use to signify and reconstruct the world around them. This exploitation of the environment has a strong sensory component, involving all their senses [5]. The way children represent things show how they understand them, and it evolves with time as they become more aware of the world around them.

At the age of four to five children's drawings are an excellent method to capture the first tangible records of a child's thinking process; however children's drawings need to be interpreted quite carefully.

Recently there have been a few instances of work exploring drawing intervention as a method to evaluate children's experience with technology [8,9,10,11,12]. In a recent work drawings coded by different researchers conveyed information about children's experience. The method used children's drawings to assess the goal fit and fun of the technology, gathering information related to user experience. Both components are commonly discussed usability factors [12]. In a previous study of ours, comparing the learning benefits of a tangible versus a graphical interface for pre-school, children's drawings were one of the methods used to assess their degree of involvement with the interfaces [8,9]. The drawings from the children that interacted with the tangible interface were very complete and detailed suggesting that the children felt actively involved, had a more engaging experience, and that the experience with the tangible interface left a greater impression on them than the game with the traditional graphical interface.

Usability evaluation through drawings?

Researchers are still trying to determine effective methods to gather usability data from children (especially within this age group) and how it can contribute to the overall design of technology for children [1,7,13].

The work presented here investigates drawings as an assessment tool for the usability of a children's interface. We devised a test where children played a computer game and for one group the game presented quite an obvious usability issue, the control froze randomly. While in this scenario other methods could have been employed to easily track that problem - just observing the frustration of the children that played

with the broken version would suffice - we wanted this test to be an easy first rehearsal for using drawings as an evaluation method. Having such a strong negative difference between the two versions of the game we would have expected it to show on children's drawings. We were wrong!

The interface

The computer game consisted of a spacecraft that children could move by handling the mouse (fig. 1). The goal of the game was to collect a series of elements that moved across the screen. To get hold of these children had to move the spacecraft with the mouse over each one. The game ended when all the elements were collected. In the final screen the ship would head to a planet with a rain of stars. In the test group the mouse would stop working during some seconds at random points in time, interrupting the flow of the game.



figure 1. The spacecraft (white and green) and the elements.

The tests

The tests were conducted at two different days with two groups of children. The teachers agreed that the tests were conducted during their normal activities. The first

test was carried with 13 children (ages four), and the second with a group of 14 children (ages five). They were carried during the morning, when all the children sat at their tables and were involved in some tasks proposed by the teachers, so that there was a quiet atmosphere.

The computer was placed at a corner of the room and one child at a time was invited to play the game. The children agreed immediately and were pleased to do it. Each child was assigned alternately the regular and the broken version. After playing the game, each child was asked if she/he would agree to draw what she/he had played (fig. 2). All the children were willing to do it. In Portugal, this is a common practice at pre-school where children are asked to draw after certain activities. The methodology allows teachers to assess children's involvement with the activity.



figure 2. A child playing the game and drawing it.

Each child played the game and made the drawing at the reading corner, one at a time, thus avoiding mutual influences, that would eventually alter the results of the drawings, since children are easily influenced by each other. After finishing the drawing the children joined the other in the class. The children that interacted with the regular version of the game were enthusiastic about it and wanted

to play it more than once. In contrast, the children that interacted with the broken version felt uncomfortable, complaining that the game would not work, and most of them eventually gave up the game, feeling really frustrated. When debriefing the children after the test, they were informed that the game had a problem and had the opportunity to play the regular game.

Evaluation of the drawings

Having collected the drawings from both groups we intended to determine to what extent it was possible to classify if the children had interacted with the regular game or the broken version (fig. 3).



figure 3. Drawings: interaction with the broken game version (left), interaction with the regular game version (right).

A group of three evaluators reviewed the drawings. Two of them were experts in the field of HCI; the third one was an expert in analyzing children's drawings. Although they were involved in the study none of them was present at the pre-school or had any contact with the drawings before.

The evaluators were asked to analyze each drawing and determine if it was done after interacting with the good or the bad game version. The evaluation was conducted individually without access to each other results.

Evaluation parameters

To classify the drawings the evaluators used the following scoring mechanism: number of drawn elements, detail and colorfulness. They considered that the drawings of the regular version would have a greater number of drawn elements, would be more detailed and more colorful, reflecting a higher involvement in the task they were performing [8,9].

Results

The results of the evaluations (table 1) showed that there was no agreement between the evaluators.

Session	Evaluator 1	Evaluator 2	Evaluator 3	Agreement between raters
1	5=38.5%	5=38.5%	8=61.5 %	46.10%
2	5=35.7%	8=57%	7=50%	50%

table 1. Number, percentage of correct evaluated drawings and agreement between the evaluators.

In the first session evaluator 1 and 2 scored 38.5%, and evaluator 3 scored 61.5% of correct results. In the second evaluation evaluator 1 scored 35.7%, evaluator 2 scored 57% and evaluator 3 scored 50% of correct results. The level of agreement between the evaluators was 46.10% in the first and 50% in the second session. This agreement percentage cannot be considered as trustworthy, since only percentages of agreement between 75% and 80% are considered relevant [2,6]. Some authors only consider a percentage of 85% as an indicator of reliability [4]. In both groups there were lively, colorful drawings, with many elements as well as drawings with just one color and representing just one or two elements. We were therefore not able to

discriminate based on the classification parameters used (number of drawn elements, detail and color) whether the child had interacted with the regular or the broken game version.

Discussion

The overall goal of this experiment was to further investigate the use of drawings as a usability evaluation methodology. In a previous study we used drawings in combination with two other methodologies: questionnaires and interviews [9]. Yet this study followed a different approach, our goal was to investigate whether drawings alone could provide any indicators on the usability of this particular interface. We tackled a quite obvious interaction problem regarding the mouse that resulted in one group having a more positive experience than the other. Definitely the two groups were exposed to two different experiences, as we could tell by watching their interaction - the complaints and frustration of the children playing with the broken game and the fact that most of those gave up. Nevertheless with our methodology we were not able to discriminate between both experiences based on the drawings alone.

Conclusion

Analyzing drawings for the purpose of collecting data is quite complex. Additionally a number of underlying factors such as environmental aspects, the presence of adults or the wish to please, can influence children's drawings. Despite these difficulties it has been shown in previous work that drawings seem to convey meaning about the fun and level of engagement with an interface [8,9,10,11,12]. Especially if they are used with other methods (interviews, observation, questionnaires, recording of children's comments while

they are drawing) they can add valuable data to usability testing. In this study however, when analyzing drawings from experiences that only differ in respect to the interface usability, there seems not to be sufficient information in the drawings to derive any conclusions about the game usability. This is clear in the lack of agreement between the raters. The results highlight the difficulty to score the drawings even when an apparently set of quantitative parameters were used.

From this we conclude that in this particular study and using the presented methodology we were not able to distinguish between groups who used a normally working game and one that did not work. Further research is still needed in this area.

Future work

As far as we know there are not known theoretical frameworks for the evaluation of drawings for usability purposes. In future work we will continue to investigate the validity of using drawings as an evaluation methodology to be used with children within this age group (e.g. finding whether drawings can be used to capture indicators of individual aspects of usability).

Acknowledgements

I would like to thank my supervisors Prof. Pedro Branco from the University of Minho, Dep. of Information Systems and Prof. Clara Coutinho from the University of Minho, Institute of Education and Psychology for their support and supervision of this work. Thank you to Prof. Eduarda Coquet from the University of Minho, Institute of Child Studies for her precious help concerning the evaluation of children's drawings.

Special thanks to the reviewers, who helped with valuable comments and suggestions to improve this paper.

Thanks to all the children who collaborated in this study and the pre-school teachers for their precious and indispensable collaboration. Thank you very much!

References

- [1] Airey S, Plowman L, Connolly D, Luckin R. Rating children's enjoyment of toys, games and media. In *3rd World Congress of the International Toy Research Association on Toys, Games and Media*, London (2002).
- [2] Cohen, J. A coefficient of agreement for nominal scales. In *Educational and Psychological Measurement*, (1960) **2**, 37-46.
- [3] Coquet, M.E. *A Narrativa Gráfica uma Estratégia de Comunicação de Crianças e de Adultos*. University of Minho, Braga, 2000.
- [4] Krippendorff, K. *Content Analysis: An Introduction to Its Methodology*. V. 5 The Sage Series, SAGE Publications, Newbury Park, 1980.
- [5] Lowenfeld, V., Brittain, W. *Creative and Mental Growth*. 8th ed. Macmillan, NY, USA, 1987.
- [6] Moore, G. W. *Developing and Evaluating Educational Research*. HarperCollins Publishers, NY, USA, 1983.
- [7] Read, J.C., MacFarlane, S.J. and Casey, C. Endurability, Engagement and Expectations: Measuring Children's Fun. In *Proc. IDC Workshop 2002*, M.M. Bekker, P. Markopoulos, M. Kersten-Tsikalkina (eds), Shaker Publishing (2002), Eindhoven, The Netherlands 189-198.
- [8] Sylla, C., Branco, P., Coutinho, C., Coquet, M.E. Storytelling through drawings: Evaluating Tangible Interfaces for Children. In *Proc. CHI 2009*. ACM Press, (2009), 3461-3466.
- [9] Sylla, C. *TUIs vs. GUIs: Comparing the Learning Benefits for Pre-school Children*, Master's Thesis, University of Minho, Guimarães, 2009.
- [10] Xu, D., Mazzone, E., MacFarlane, S. In search for evaluation methods for children's tangible technology. In *Proc. IDC 2006*. ACM Press (2006), 171-172.
- [11] Xu, D., Read, J.C., Sheehan, R. In Search of Tangible Magic. In *22nd BCS British- HCI 2008*. Liverpool, UK.
- [12] Xu, D., Read, J.C., Sim, G., McManus, B. Experience it, draw it, rate it: capture children's experiences with their drawings. In *Proc. IDC 2009*. ACM Press (2009), 266-270.
- [13] Zaman, B., Abeele, V. How to Measure the Likeability of Tangible Interaction with Preschoolers. In *Proc. CHI Nederland (2007)*, 57-59.