Adapting the Heuristic Evaluation Method for Use with Children

Stuart MacFarlane **Anastasia Pasiali**
Child-Computer Interaction Group
Department of Computing
University of Central Lancashire
Preston, Lancashire, UK
sjmacfarlane@uclan.ac.uk apasiali@uclan.ac.uk

ABSTRACT
Heuristic Evaluation is a method that relies on experts (normally in usability) to predict problems that users will have with a system, by comparing the system against a set of heuristics. The paper describes problems with the use of the method with software designed for children, in particular the problem that it is hard for (adult) experts to put themselves in the position of child users. A variation of the method wherein children themselves carry out a simplified form of heuristic evaluation is proposed, and initial observations from its use are reported.

Keywords
Children, evaluation, heuristics.

INTRODUCTION
A heuristic evaluation method is an analytical evaluation method that relies on comparing an interface (or an interface design representation) against a pre-defined set of heuristics. The aim of such a comparison is to find problematic aspects of the design in order to improve them during the follow-up stages of the design process. The 'heuristics' are generic rules or design guidelines or principles for good interaction design.

Traditionally, the term 'heuristic evaluation' is used to describe the method as it was introduced by Nielsen and Molich [5] for use with their own set of usability guidelines. In this paper, we shall use the term also to refer to evaluation against other sets of guidelines or rules that may pertain to usability or to fun and enjoyment.

Heuristic evaluation is carried out by an 'expert' or, better still, a team of experts. An 'expert' in this context is someone trained in the use of the method, who is not a user, and preferably not a designer of the product being evaluated. The experts' expertise is normally in usability (or whatever other aspect of the product that is under evaluation); they are not necessarily domain experts.

The method can be used in a summative way with a complete product, though it was designed primarily for use formatively using a prototype.

The principle of the method is to select an appropriate set of heuristics or guidelines and examine the product or prototype to see whether, how, and how severely it breaks the guidelines. Each of the problems found in this way is likely to be a problem also for actual users.

EVALUATING CONSTRUCTS OTHER THAN USABILITY
The vast majority of the literature on heuristic evaluation is concerned with evaluations of usability. However, guidelines exist for other aspects of the user experience of interactive products, such as 'fun' [3], 'instructional design' [6], and 'game playability' [1]. These can be less successful in use, for two reasons.

Firstly, guidelines written for designers are not exactly the same as heuristics for use in evaluation. Although they describe the same constructs, they are expressed differently. They may be too specific, or too general, for use as heuristics. Often they need to be rephrased before they can be used successfully in a heuristic evaluation.

Secondly, Nielsen's usability heuristics have undergone extensive testing and several iterations of design. The other heuristics have been much less widely used and may need further work before they are ideal for use in heuristic evaluation.

PROBLEMS WHEN USING HEURISTIC EVALUATION FOR CHILDREN'S PRODUCTS
The Heuristic Evaluation method has been widely applied, using adult evaluators, in the evaluation of children's products.

The method assumes that there is a substantial, and largely agreed, set of interface design guidelines that apply to the particular situation under investigation. Where the users are adults, this is now a reasonable assumption for a wide range of situations, and this is a cost effective and useful method. Where the users are children, however, there is much less established 'theory' about interface design. Consequently, using this method when evaluating products for children is likely to be less effective than it is when used for adults' products.

In addition, the method requires the expert to identify breaches of guidelines that will cause problems for children; this requires them to have some insight into how children would behave in particular circumstances. This may not be easy for evaluators without extensive experience in working with children. Children are surprisingly resilient to issues that adults perceive as potential problems. For example, in a study reported by MacFarlane et al [2] adult evaluators consistently picked out a problem with the wording of instructions in a product for young children, and rated it as quite severe. However, in observation studies of the product being used by children, none of them showed any evidence of a problem with the instructions, nor did they ask about them. Most of them probably couldn't have understood the wording, but this was not a problem for them; they simply guessed what the instructions must have been, or asked by-standers. Children are used to managing without instructions; not being able to read and understand them isn't necessarily a problem!

HEURISTIC EVALUATION METHODS USED BY CHILDREN

It is possible to use Heuristic Evaluation methods where the evaluators are not experts but children, of or above the age range for which the product is intended. Such evaluators will necessarily be less 'expert' in usability, but have the major advantage of a better understanding of which factors are of importance to the intended users. MacFarlane et al [2] showed that children as young as 7 or 8 could reliably differentiate between constructs such as 'ease of use' and 'fun' in a user interface, so older children ought to be able, in principle, to carry out a heuristic evaluation. The work described in this paper tests this hypothesis.

A potential difficulty here is that the children will not necessarily fully understand the implications of problems in the interface, having had less experiences of interfaces in general, and of evaluation in particular.

THE EXPERIMENTAL WORK

The product chosen for evaluation was a web-based French language tutorial, freely available on the BBC website (http://www.bbc.co.uk/languages/french/lj/). The product does not specify a particular age range, but is suitable for learners of the age tested. We decided to evaluate it for both usability and fun, using heuristics based on the usability ones by Nielsen [4] (Table 1) and on the guidelines for fun by Malone [3] (Table 2).

Visibility of system status
The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.
Match between system and the real world
The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.
User control and freedom
Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.
Consistency and standards
Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.
Error prevention
Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.
Recognition rather than recall
Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.
Flexibility and efficiency of use
Accelerators -- unseen by the novice user -- may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.
Aesthetic and minimalist design
Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.
Help users recognize, diagnose, and recover from errors
Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.
Help and documentation
Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

Table 1 – usability heuristics [4]

Challenge
Not too easy, not too hard Goals should be clearly identifiable. The user should be uncertain about achieving the goal (not too easy – not too hard) Personal performance feedback should be frequent, clear, constructive and encouraging.

Fantasy
Evoke mental images of objects/situations not present. Appeal to imagination and encourage identification with characters or contexts.
Curiosity
Sensory curiosity supported by variable auditory and visual effects. Cognitive curiosity supported by surprises, paradoxes, humour, dealing with topics that already interest the user.
Control
The environment should be responsive. The user controls the outcome of the system.

Table 2 – fun guidelines [3]

The first stage was to adapt the heuristics into a form usable by younger evaluators. For the usability heuristics, all that was necessary was to simplify the language, removing the 'usability jargon', including the names for each of the heuristics. One of the heuristics was split into two for clarity. The resulting heuristics are listed in Table 3.

The fun guidelines were found to be rather broad; their use as heuristics could be made easier by splitting some of them into two or more heuristics. Additionally, the language was simplified; see Table 4.

The software keeps the user informed about what is going on
The software includes the users' language rather than words, phrases or concepts that are not familiar to the user.
Information appears in a natural and reasonable order
Users are able to exit pages and undo mistakes (for example to go back or to start again)
The system is consistent. Users do not have to wonder whether different words, situations or actions mean the same thing.
The software is designed to help users avoid making errors
The user does not have to remember how to use the software. Information required is visible or easily accessed
The software allows frequent users to use shortcuts and adjust settings to suit
The design is simple and there is no irrelevant information
Errors messages are explained in plain language, clearly indicate the problem and recommend a solution
The software provides appropriate help and documentation which is ease to access and relevant to the users' needs

Table 3 – revised usability heuristics

The goals of the software are clear to the user
The tasks are not too easy and not too hard
The user gets frequent, clear feedback that encourages him/her to carry on
The system allows the user to use his/her imagination and identify with the characters and contexts
The software includes sound and visual effects.
The system includes surprises, humour and interesting things for the user
The user feels in control of the system.

Table 4 – revised fun heuristics

The first evaluation study was carried out by 15 children aged 13-14 at a local co-educational high school. The children covered the normal ability range for their age. The work was done, with the cooperation of their teacher, in their normal Information Technology class. The purpose of evaluation, the heuristic method, and the concepts of usability and fun were briefly explained to them.

Half of the children did an evaluation of the product using the usability heuristics, and half used the fun heuristics. These were allocated alternately around the room, to reduce collusion. Initial results are encouraging; the children were almost all interested participants, they asked questions of the researcher (and of their teacher, who was present throughout). Typically they spent 20-25 minutes on the evaluation task, after spending a few minutes initially exploring the software.

Further studies are planned, to increase the amount of data and the range of ages of children. Detailed analysis of the results from this first study may indicate that some of the heuristics need rephrasing.

In addition, similar evaluations, using the same heuristics, are to be done by adult experts, to provide a basis for comparison.

More detailed results of the study will be available in time for the workshop.

REFERENCES

1. Desurvire, H., Caplan, M., and Toth, J. A. (2004). Using heuristics to evaluate the playability of games. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria, April 24 - 29, 2004). CHI '04. ACM Press, New York, NY, 1509-1512.
2. MacFarlane, S., Sim, G., and Horton, M. (2005) Assessing Usability and Fun in Educational Software, in Proceedings of Interaction Design and Children 2005, Boulder CO, USA, ACM Press.
3. Malone, T.W. Heuristics for designing enjoyable user interfaces: Lessons from computer games. In John C. Thomas and M. L. Schneider (Editors), Human Factors in Computing Systems, Norwood, NJ: Ablex Publishing Corporation, 1982.
4. Nielsen, J. (1994). Heuristic evaluation. In Nielsen, J., and Mack, R.L. (Eds.), Usability Inspection Methods, John Wiley & Sons, New York, NY.
5. Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. Paper presented at the ACM CHI'90

Conference on Human Factors in Computing Systems, Seattle, WA.

6. Quinn, C. N. (1996). Pragmatic evaluation: lessons from usability. In A. Christie, P. James & B. Vaughan (eds.) Proceedings of ASCILITE 96. Uni SA, Adelaide. p.437-444.