Using Survey Methods for Design and Evaluation in Child Computer Interaction

Janet Read
Child Computer Interaction Group
University of Central Lancashire
Preston, PR1 2HE
+44(0)1772 893285
jcread@uclan.ac.uk

Kim Fine
Department of Computing
University of Central Lancashire
Preston, PR1 2HE
bonnielll@aol.com

ABSTRACT
This paper begins with a review of some of the current literature on the use of survey methods with children. It then presents four known concerns with using survey methods and reflects on how these may impact on studies in Child Computer Interaction. The paper presents some guidelines for HCI researchers and developers intending to use survey methods in their studies with children, and closes with some thoughts about the use of survey methods in this interesting but complex area.

Keywords
Survey methods, Questionnaires, Interviews, Children, Evaluation

1. INTRODUCTION
The method of eliciting information by questioning is commonly referred to as a survey method. There are many such methods and the term survey has many meanings but for the purposes of this paper, survey methods are defined as questionnaires, rating scales and structured interviews [1]. Thus, free discussion and free form reporting is not especially considered.

Surveys are a long established instrument for data gathering and as early as the 1890's they have been reported as being used with children [2]. However, research about the efficacy of the different methods of surveying children is relatively scarce and in particular, when children are asked to contribute opinions, as is the case in self-reported research, studies that examine the validity and reliability of the children's responses are rare [3]

2. WHY ASK CHILDREN?
In the field of Child Computer Interaction [Read, 2005 #726] it is common to find studies that report the use of survey methods with children. In some of these studies (loosely referred to hereafter as design), children are asked to contribute ideas and suggestions for future or partially completed designs. Examples include the use of surveys to elicit detail about the mental models that children have [4], or their use to gather requirements for interfaces [5]. More commonly, surveys are used in evaluation studies, where children are asked to comment on the appeal or usefulness of a product or supply some sort of product rating [6].

It is sometimes useful to separate these two instances of use, particularly where methods that are used in surveying are dependent on context, however, in all instances, to different degrees, a survey method is a gathering of opinions and so in this paper, the authors prefer to remain open minded and consider the issues that effect both attitudinal and factual data gathering by considering survey methods with a single treatise.

There are several valid reasons for asking children for their opinions of interactive products. One is that adults and children live in different worlds and for that reason adults may not understand what children want, "*Survey researchers are realising that information on children's opinions, attitudes and behaviour should be collected directly from the children; proxy-reporting is no longer considered good enough*." [7]. Secondly, there is a move to include children in decisions about their own environments; this has arisen from a greater awareness that children are actors and participants rather than onlookers in society. "*In most of the western world, it is now recognised that children have a voice that should be heard and there is a new demand for research that focuses on children as actors in their own right*." [8]. A third reason for talking to children about their interactive technologies, perhaps for some people the most motivating, is that involving children in the design and evaluation of their own artefacts is fun and rewarding for researchers, developers and, more importantly, for children [9].

3. WHAT CAN GO WRONG?
Surveys methods rely on the use of a question –answer process. This process can be vitally important in contributing to the success of a survey in eliciting reliable data. There are four stages in a question-answer process:

1. Understanding and interpreting the question being asked.
2. Retrieving the relevant information from memory.
3. Integrating this information into a summarised judgement.
4. Reporting this judgement by translating it to the format of the presented response scale.[10]

Researchers often discuss the importance of the question-answer process in determining the reliability of responses provided by children in surveys [3]. Factors that impact on question answering include developmental effects; language, reading age, and motor abilities, as well as temperamental effects including confidence, self-belief and desire to please.

Research into the completion of surveys has revealed four major issues that are particularly important in understanding how children respond to surveys. The first two issues are phenomena that are partly temperamental and partly cognitive in nature; these will have an impact on the design of survey studies. The second

two are more to do with language and age and are rather more concerned with the detailed design of the question and answer processes.

3.1 Satisficing and Optimising

Satisficing theory identifies two processes that explain some of the differences in the reliability of responses, especially in surveys where respondents are being asked to pass attitudinal judgments [11]. For research validity, optimising is the preferred process; this occurs when a survey respondent goes thoughtfully and carefully through all four stages of the question and answer sequence. Satisficing is the opposite approach and occurs when a respondent gives more or less superficial responses that generally appear reasonable or acceptable, but without having gone through all the steps involved in the question-answer process.

The degree or level of satisficing is known to be related to the motivation of the respondent, the difficulties of the task, and the cognitive abilities of the respondent [7]. It appears obvious therefore, that if a child misunderstands a question or finds it difficult to answer then the child is susceptible to 'satisfice'.

3.2 Suggestibility

Suggestibility is particularly important with relation to survey research with children, because it "*concerns the degree to which children's encoding, storage, retrieval and reporting of events can be influenced by a range of social and psychological factors*." [12]. In any survey, the interviewer or researcher has an effect. Even when the interviewer is trying hard not to impact on the question answer process, when the respondents are children it is sometimes impossible to not intervene. In one study it was reported "*there were many silences that needed some input if only to make the children less uncomfortable*." [5].

Even where there is no deliberate intervention the interviewer has an effect. In one study it was shown that children are likely to give different responses depending on the status of the interviewer. This was illustrated when a research assistant pretending to be a police office asked children questions about their experience with a babysitter. The children then assumed that the nature of the experience was bad and thus the interviews yielded inaccurate and misleading results [13]. It seems that authority figures may inevitably yield different results, as the child may want to please the person administering the survey [14].

The gender and age of the interviewer or person conducting the survey can also have an effect on the reliability or detail of responses provided by children. Borgers et al, (2004) discuss an example stating:: "*There is anecdotal evidence from surveys on drugs in Germany that teenagers were far more open to elderly female interviewers and not to the young or youngish interviewer*s." [3].

3.3 Specific Question Formats

The way in which children are asked questions in surveys has an impact on the reliability of the response. Breakwell et al, (1995) report that "*There is a strong acquiescence response bias in children: children tend to say 'yes', irrespective of the question or what they think about it.*" [10]. In one study with 5-year-old children there were several inaccuracies in questions that relied on the yes/no format [14].

Free-recall questions have been shown to be useful with children, especially in spoken surveys. One study involved children who had experience of being treated in an emergency room for an injury. A few days later, children were interviewed with free recall question formats such as "Tell me what happened" and specific questions like "Where did you hurt yourself?" both being used. It was shown that as the questions became more specific i.e. "Did you hurt your knee?" the response reliability decreased [15].

One widely used question format is the use of Visual Analogue Scales (VAS). A VAS uses pictorial representations that children use to identify their feelings or opinions. This approach has been adopted as an alternative to the traditional open-ended and closed question formats although some researchers suggest that VAS can only be used with children aged around seven and over [16]. Studies in Child Computer Interaction have shown them to be useful for younger children, but have also noted that when these scales are used to elicit opinions about software or hardware products, younger children are inclined to almost always indicate the highest score on the scale [17].

Below are two examples of Visual Analogue Scales developed for children for different purposes.

Figure 1: Wong-Baker pain rating scale. [18]

Figure 2: Smileyometer scale for eliciting children's opinions [19].

3.4 Language Effects

Children have varying abilities in spoken and written language and this makes the design of questions for surveys problematic. Research suggests that language in surveys is especially important and that vague and ambiguous words should be avoided [7]. With visual analogue scales, or with multi-choice responses, the response options should be completely labelled to help children to produce more reliable responses [20].

Children are known to take things literally; in one study it was noted that when a group of children were asked if they had been on a 'school field trip' they replied 'no' because they did not refer to the trip as a 'school field trip.' [21]. In a more recent study, it was noted that when children were asked how good they thought a writing activity had been, some children gave an opinion of their writing as a product, thus interpreting the question in a completely unexpected way [22].

4. IMPACTS ON STUDIES IN CCI

It is inevitable that these four factors will affect response quality and reliability when using surveys with children. Some of these effects will be more pronounced with younger children than with teenagers. In addition, it has been noted that in some instances, children's responses are not very stable over time so it may be that all that can be elicited from a survey is a general feel for a product or a concept with a particular group of children at a particular time [23].

In studies in Child Computer Interaction, the stability of responses and the reliability of responses are generally not critical (as could be the case where a child is being interviewed as part of a criminal investigation). Given this, there are several useful approaches that can be taken to make the surveying process valuable and satisfactory for all the parties.

1. Keep it short

Whatever the children are asked to do, make it fit their time span. This will reduce the effect of satisficing by keeping their motivation high. For young children, five minutes spent in a written survey is generally long enough, more time can be given, as the children get older.

2. Pilot the language

In a survey using written language, children will take short cuts if they cannot read the questions. Teaches can be useful in checking to see if the words used in the survey make sense, they may point out where words may mean something different to children. Avoid ambiguity by piloting with sample children.

3. Provide assistance for non / poor readers

Even with the language checked, there will be some children who may understand the words but not the questions. Try to read out written questions if possible, doing this for all the children (as some will not admit to not understanding the questions).

4. Limit the writing

Children often do not write what they want to say, as they cannot spell the words they want, cannot find the words for things they want to say, or cannot form the letters for the words that they have in mind. Children can be helped by encouraging the drawing of pictures, the use of images and by providing essential words for them to copy.

5. Use appropriate tools and methods

Reduce the effects of suggestibility and satisficing by using special methods. The Fun Toolkit provides tools to assist children in discriminating between rival products [24]. In interviews, use visual props to help articulate ideas. If interviewing, consider taping the discussion so that the amount of 'suggesting' can be examined later.

6. Make it fun

Introduce glue, scissors, sticky tape or coloured pencils to make the experience fun for the children. If at all possible print questions in colour and supply thank you certificates when the children have finished participating

7. Expect the unexpected

Have a back up plan. If an entire project depends on the results of a survey with children it may well fail! Triangulate where possible by gathering self reported data, observational data and some post hoc thoughts from researchers and participants.

8. Don't take it too seriously

One of the great pitfalls in research and development work is to read too much into data. The information gained from a single group of children in a single place is not likely to be especially generalisable. Avoid the temptation to apply statistical tests to children's responses, rather look for trends and outliers!

9. Be nice

As outlined earlier, interviewer effects are significant. To get the most from children, interviewers and researchers need to earn the right to talk to them. This may require several visits and may require an investment of time to learn about their culture and their concerns.

There is no doubt that designing and carrying out good surveys takes practise and patience but following these guidelines may avoid many of the common errors and minimise harmful effects. .

5. CONCLUSION

Because a survey is, by definition, designed, it will always be restrictive. Researchers and developers of interactive products are generally not specialists in survey design and so invariably produce questions and suggested answers that are far from perfect. It is common, and not unexpected, to find that in many studies, the questions are asked in such a way that the answers are invariably the ones the survey designers wanted to hear.

Given the inherent difficulties with survey methods for children, and the survey designer's inadequacy, a case could be made for discouraging these methods in Child Computer Interaction. This approach might gain favour with the empiricists but the value of the survey method to the Child Computer Interaction community is not its validity or its generalisability, but rather the opportunity that these methods provide for researchers and designers to interact with children, to gather their language, and to value their differences. A survey can say much more than $p < 0.05$; it can say that some children find things so good they tick the box umpteen times, it can say, '*I don't want to pick between these things, they were all good*', it can say *my favourite colour is blue even if you never asked me*, and it can say, *me and my friend filled this in together*.

Perhaps success in a survey in Child Computer Interaction is not to do with stability of responses or reliability of reports but is measured by the answers to two questions for the survey designer, these being: 'Did I learn anything useful? Did I do anything useful?

It is a privilege to be able to carry out design and evaluation surveys with children. Researchers and developers get to see into the children's worlds and get to glimpse at their dreams and ideals. This requires care and concern, in the words of WB Yeats, "*Tread softly because (you) we tread on (my) their dreams*". It is especially important to neither waste the children's time nor ride roughshod over their opinions.

The guidelines presented in this paper are intended to assist practitioners to carry out careful and gently executed surveys that respect the children and protect their ideals.

Much of the literature pertaining to surveying children focuses on what not to do and on the precautions that need to be taken to safeguard the data, future research by the authors will focus on the cost / benefits of surveys in CCI and on refining the methods that need to be taken to provide a special experience for the children.

Acknowledgements

Hundreds of children have freely shared their opinions and views with the authors. They are acknowledged here for their willingness to provide windows into their worlds, for moments of amusement, and for the many special insights they have provided.

6. REFERENCES

1.Greig, A. and A. Taylor, *Doing Research with Children*. 1999, London: Sage.

2.Bogdan, R.C. and S.K. Biklen, *Qualitative research for education: An introduction to theory and methods*. 1st ed. 1998, Boston: Allyn and Bacon.

3.Borgers, N., J. Hox, and D. Sikkel, *Response Effects in Surveys on Children and Adolescents: The Effect of Number of Response Options, Negative Wording, and Neutral Mid-Point.* Quality and Quantity, 2004. **38**(1): p. 17 - 33.

4.Read, J.C., S.J. MacFarlane, and C. Casey. *What's going on? Discovering what Children understand about Handwriting Recognition Interfaces*. in *IDC 2003*. 2003. Preston, UK: ACM Press.

5.Read, J.C., S.J. MacFarlane, and A.G. Gregory. *Requirements for the Design of a Handwriting Recognition Based Writing interface for Children*. in *Interaction Design and Children 2004*. 2004. Maryland, US: ACM Press.

6.Read, J.C., S.J. MacFarlane, and C. Casey. *Measuring the Usability of Text Input Methods for Children*. in *HCI2001*. 2001. Lille, France: Springer Verlag.

7.Borgers, N. and J. Hox, *Item Non response in Questionnaire Research with Children.* Journal of Official Statistics, 2001. **17**(2): p. 321 - 335.

8.Borgers, N., E.D. Leeuw, and J. Hox, *Children as Respondents in Survey Research: Cognitive Development and Response Quality*. Bulletin de Methodologie Sociologique, 2000. **66**: p. 60 - 75.

9.Read, J.C., *The ABC of CCI*. Interfaces, 2005. **62**: p. 8 - 9.

10.Breakwell, G., *Research Methods in Psychology*. 1995, London: SAGE Publications.

11.Krosnick, J.A., *Response Strategies for coping with the Cognitive demands of attitude measures in surveys.* Applied Cognitive Psychology, 1991. **5**: p. 213 - 236.

12.Scullin, M.H. and S.J. Ceci, *A Suggestibility Scale for Children.* Personality and Individual Differences, 2001. **30**: p. 843 - 856.

13.Tobey, A. and G. Goodman, *Children's eyewitness memory: effects of participation and forensic context.* Child Abuse and Neglect, 1992. **16**: p. 807 - 821.

14.Bruck, M., S.J. Ceci, and L. Melnyk, *External and Internal Sources of Variation in the Creation of False Reports in Children.* Learning and Individual Differences, 1997. **9**(4): p. 269 - 316.

15.Peterson, C. and M. Bell, *Children's memory for traumatic injury.* Child Development, 1996. **67**(6): p. 3045 - 70.

16.Shields, B.J., T.M. Palermo, J.D. Powers, S.D. Grewe, and G.A. Smith, *Predictors of a child's ability to use a visual analogue scale.* Child: Care, Health and Development, 2003. **29**(4): p. 281 - 290.

17.Read, J.C., P. Gregory, S.J. MacFarlane, B. McManus, P. Gray, and R. Patel. *An Investigation of Participatory Design with Children - Informant, Balanced and Facilitated Design*. in *Interaction Design and Children*. 2002. Eindhoven: Shaker Publishing.

18.Airey, S., L. Plowman, D. Connolly, and R. Luckin. *Rating Children's Enjoyment of Toys, Games and Media*. in *3rd World Congress of International Toy Research on Toys, Games and Media*. 2002. London.

19.Read, J.C., S.J. MacFarlane, and C. Casey. *Expectations and Endurability - Measuring Fun*. in *Computers and Fun 4*. 2001. York, England.

20.Borgers, N., J. Hox, and D. Sikkel, *Response Quality in Research with Children and Adolescents: The Effect of Labelled Response Opinions and Vague Quantifiers.* International Journal of Public Opinion Research, 2002. **15**(1): p. 83 - 94.

21.Holoday, B. and A. Turner-Henson, *Response effects in surveys with school-age children.* Nursing Research (methodological corner), 1989. **38**(248 - 250).

22.Read, J.C., S.J. MacFarlane, and M. Horton. *The Usability of Handwriting Recognition for Writing in the Primary Classroom*. in *HCI 2004*. 2004. Leeds, UK: Springer Verlag.

23.Vaillancourt, P.M., *Stability of children's survey responses.* Public opinion quarterly, 1973. **37**: p. 373 - 387.

24.Read, J.C., S.J. MacFarlane, and C. Casey. *Endurability, Engagement and Expectations: Measuring Children's Fun*. in *Interaction Design and Children*. 2002. Eindhoven: Shaker Publishing.