Considerations and Methods for Usability Testing with Children

Malene Hjortboe Andersen, Md. Saifuddin Khalid(✉), and Eva Irene Brooks

Department and Learning and Philosophy, Faculty of Humanities, Aalborg University, Kroghstræde 3, 9220 Aalborg, Denmark
malenehjortboe@gmail.com, professorkhalid@gmail.com, eb@learning.aau.dk

Abstract. In this paper, the authors draw on methods used in the field of interaction design, emphasizing a user-centred design approach including methods such as usability testing, design metaphors, interview with users, video observations, focus groups, and think aloud sessions. However, a challenge of these methods is that they are designed for adults and are not necessarily appropriate to investigations including children. The guiding questions for this systematic literature review are (1) the motivation for conducting usability tests with children, and (2) the kind of methodological, practical, and ethical considerations that should be considered when involving children in usability studies. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist and PRISMA flow diagram are applied in order to assure the quality of the process of this investigation. Nine articles are analyzed and then synthesized by applying the constant comparative method. The synthesis of the literature review is based on the identified thematic priorities, which are categorized as follows: (1) the motivation for involving children as test persons in design processes, (2) definitions of usability, (3) practical considerations, (4) methodological considerations, and (5) ethical considerations.

Keywords: Child-computer interaction · Usability testing · Usability evaluation · Usability testing methods

1 Introduction

In this paper, the authors will draw on methods used in the field of interaction design [1], emphasizing a user-centered design approach including methods such as usability testing, design metaphors, interview with users, video observations, focus groups, and think aloud sessions (c.f. [2, 3]).However, a challenge of these methods is that they are designed for adults and are not necessarily appropriate to investigations including children [4]. The needs, skills, terminologies, and desires of children are essentially different from those of adults (cf. [5]). In their studies, Read and Markopoulos [6] and Markopoulos and Bekker [7] highlight that current literature seldom focuses on children. At the same time, research, as well as industry, target children as a user group or as consumers, resulting in an increased need to consider the context and inquiries that are

A.L. Brooks and E. Brooks (Eds.): ArtsIT/DLI 2016, LNICST 196, pp. 228–238, 2017.
DOI: 10.1007/978-3-319-55834-9_27

unique to the field of child-computer interaction. This paper will try to meet some of these challenges.

When including children in usability studies, Druin [5] differs between four roles the children can employ: user, tester, informant, and design partner. The main difference is the distribution of power between the children and the researchers. The first two terms, Druin [8] defines as reactive participants, including methods such as children observing other children [5], play sessions [9], and post-task interviews [22]. The last two terms, informant and design partner, Druin [8] classifies as participative users. Here, the design inquiries include methods such as cooperative low-tech prototyping [8], drawings [10], and technology immersion [5].

The overarching goal of the research was to establish an overview of usability studies carried out including children. For the literature review, the guiding questions related to (1) the motivation for conducting usability tests with children, and (2) the kind of methodological, practical, and ethical considerations that should be considered when involving children in usability studies.

2 Methods

In this study, Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist and PRISMA flow diagram [11] are applied in order to assure the quality of the process of this investigation. The checklist is applied as guidance to ensure that the study includes the central areas of the systematic literature review. The flow diagram is used to document the study searches, the search results, and inclusion and exclusion criteria.

The searches are conducted in the Educational Resources Information Center (ERIC) and Scopus databases. Different combinations of the following keywords and synonyms are applied: *usability testing*, *usability evaluation*, *children*, *interaction design*, *methods*, and *guidelines*. It is identified that the keyword *children* returns more relevant results than the use of *school students* or *primary school.* The searches are restricted to only English and peer-reviewed articles.

Figure 1 illustrates the research methodology, including searches, screenings, as well as inclusion and exclusion process. The figure is based on the PRISMA flow diagram [11], which includes four stages: identification, screening, eligibility, and included. In the *identification* phase, the search in the databases resulted in a total of 105 articles. During the *screening*, 80 articles were excluded while studying the titles, abstracts, and keywords of the 105 articles. The exclusion criteria were as follows: an article is excluded if (1) the article is not about children aged 5–17 years of age, or (2) the primary focus of the article is on usability tests of a specific product, and not on the methods or considerations focusing on the involvement of children in the testing process. In the *eligibility* phase, the full-text of each of the 25 articles was read, and 16 articles were excluded by applying the two exclusion criteria. Also, it is noteworthy that 5 of the 16 excluded articles were written by the same authors, and the content of the articles was similar. Ideally, those five articles should have been included as they cannot be rejected according to the rejection rule. Eventually, a total

of 9 articles has been included in the qualitative analysis and synthesis. The selected nine articles are published between 1997 and 2015.

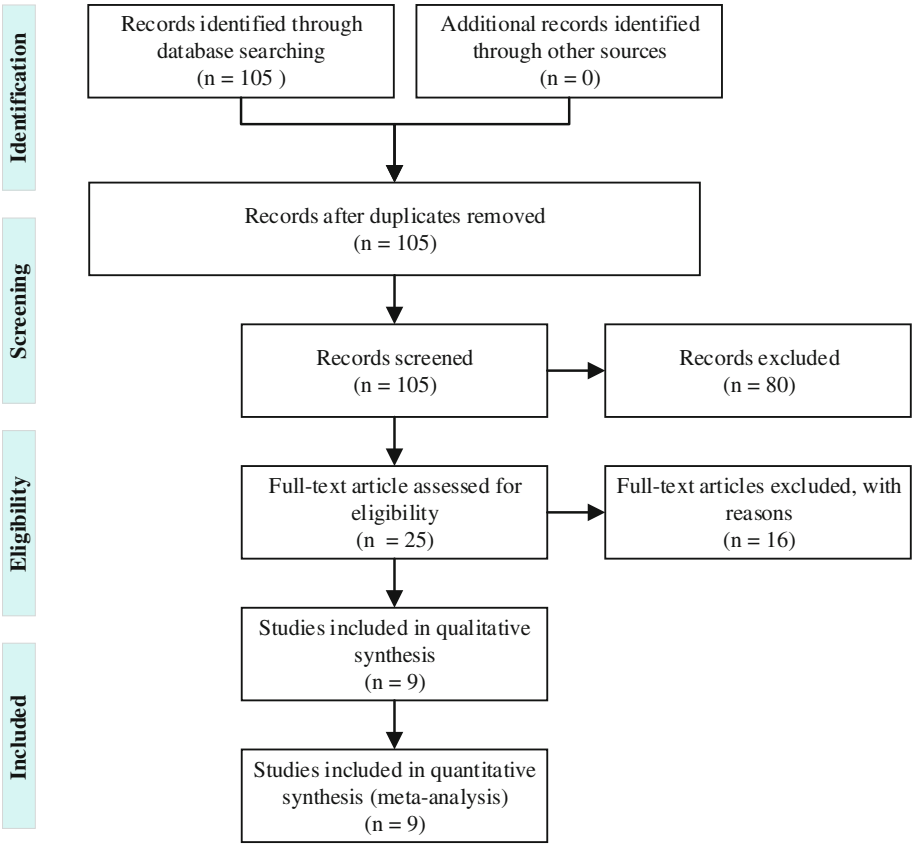

Fig. 1. PRISMA flow diagram [11]

The nine articles are analyzed and then synthesized by applying "the constant comparative method" [12]. The method is applied to create an overview of the key themes that appear in the selected literature. It is an iterative process, where the relevant themes of the individual articles are identified, and used to analyze the subsequent articles. The process is as follows: the first article is read, and relevant themes highlighted and, then, the next article is read applying the same themes and identifying more themes. A table is made to record and summarize the themes identified in each of the articles (see Table 1). If one or more of the themes in an article matches with previously recorded themes, then a check mark is placed on the table. If a subsequent article's theme does not match with any of the existing themes, then a new theme is added. This process is repeated until the analysis does not result in any relevant new themes.

Table 1. Overview of articles and identified themes

Publication year	1997	2002	2003	2005	2008	2012	2013	2014	2015
Authors	Hanna, L. Risden, K., and Alexander, K.J.	Druin, A.	Markopoulos, P., Bekker, M.	Als, B.S., Jensen, J.J., Skov, M.B	Read, J.C., Markopoulos, P.	Rounding, K.,Tee,K., Wu, X.,Guo, C.,Tse, E.	Khanum, M.A., Trivedi, M.C	Read, J.C., Fitton, D., Horton, M.	Read, J.
Country	USA	USA	Netherlands	Denmark	England	England	India	England	England
Article type & Themes	J	J	J	C	C	C	C	C	J
The motivation to do usability testing with children	X	X			X	X	X		
Necessity of child-friendly environment	X	X		X					
The duration of the test should be considered	X		X	X					X
Children will be recruited from the age group	X	X		X					
The child must sit with an adult during the test	X	X					X		
Definition of usability			X	X			X		
Comparison of think aloud and constant comparative			X	X			X		
Children should be recruited based on language skills					X			X	X
Children must cooperate in pairs					X	X		X	X
Children should know the purpose of the test					X			X	X
Children should be thanked for participating in the test					X			X	X
Natural surroundings give the best results							X		X

3 Qualitative Analysis and Synthesis

The synthesis of the literature review is based on the identified thematic priorities (see Table 1), which are categorized as follows: (1) the motivation for involving children as test persons in design processes, (2) definitions of usability, (3) practical considerations, (4) methodological considerations, and (5) ethical considerations.

3.1 The Motivation to Involve Children as Testers in Design Processes

The child's role as a test person was first discussed in the literature of the late 1980s and early 1990s by Allison Druin [13]. The discussion was particularly relevant in

connection Seymour Papert's research group, who developed the *Logo* programming language: "It may well have been Papert and his colleagues' deeply held belief in children as builders, scientists, and learners that led to the early inclusion of children in the technology design process" [5]. The research group placed great emphasis on the involvement of children as test persons in the design processes. This resulted in a discovery of numerous problems, which adult test persons would never have been able to address [5]. Several large companies followed and since then it became common for manufacturers and designers to recruit children for testing [5].

The motivation for including children in the design process has increased as children have been savvy users of digital devices such as computers, phones, tablets and game consoles [5, 14]. "*Children are fast becoming tomorrow's power-users of everything from the Internet to multimedia authoring tools*" [5]. In this context, it has been essential to address how children can be included in the design of new technologies and at the same time which differences there is between involving children and adults as test persons.

Research within the child computer interaction community (CCI) has shown that there is an immense difference between the evaluations of designs tested by adults and the ones including children [7]. Children can be considered as a separate group with its culture, norms and complexities [5]. "*Children are not miniature adults, but they have their own set of preferences, perception, style, likes, and dislikes. When designing technology for children their preferences should be taken into account.*" [14] Many designers often have their assumption about children or include personal experience of being a child [5]. These personal assumptions are not enough to represent today's children: "*Children are very different from adults — they have developing motor skills, limited reach, short attention spans, limited exposure to traditional user interfaces and social protocols."* [13] One central motivation to include children in a design process is that "*Children are extremely honest in their feedback and comments concerning technology […]"* [5].

Druin identified "four main roles that children can play in the technology design process: *user*, *tester*, the *informant* and *design partner*" [5]. Each of the four roles are based on three underlying dimensions: (1) relationship to adults (which are indirect, including feedback, dialogue, and elaboration); (2) relationship to technology (which concern ideas, prototypes and products); and (3) goals for the inquiry (which include developing educational theory, questioning impact of technology, and improved usability/design) [5]. The motivation of recent child interaction design literature contributes to the understanding and the application of the four roles of children and the three dimensions in the design process.

3.2 Definitions of Usability

Usability testing is one of the disciplines that takes into account the preferences of children [7, 14]. Usability testing has been studied and used for many years, and the overall aim is to "*identify some of the key interaction problems in user interfaces*" [15] Today, it is common that users, first of all, check how easy it is to understand product functionalities. This indicates that users have become more aware of the products' usability [14]. "*Usability is most*

often defined as the ease of use and acceptability of a system for a particular class of users carrying out specific tasks in a specific environment. Ease of use affects the user's performance and their satisfaction while acceptability affects whether the product is used" [14, 16].

Usability testing can take the form of many different methods; the most common is a so-called think-aloud test [7]. According to Khanum and Trivedi [14], it is best to conduct usability testing in the early design stages, since it is central to disclose any problems or so-called 'bugs' early.

In usability tests with children, Markopoulos and Bekker [7] address the lack of a common definition of the term usability, and states: "*At present we are conducting various studies to determine appropriate evaluation criteria for interactive products for children, as how to test not only for how usable a product is found but, also, how much fun children experience while interacting with it*" [14]. Markopoulos and Bekker [7] also points out that there can be substantial differences in the purposes of the tests, depending on whether test persons are adults or children. The next section will focus on this adaptation.

3.3 Considerations

There is an extensive difference between testing with children and testing with adults: "*Working with children to evaluate interfaces requires that some adaptations be made to techniques traditionally used with adults*" [13]. Many of the selected articles are concerned with how this adaptation can take place and make more suggestions about considerations the researcher should be informed beforehand. "*There are practical concerns around arranging studies and recruiting children, there are methodological concerns in terms of ensuring that children can contribute in meaningful ways, and there are ethical concerns around the meaning of the children's participation.*" [17] Based on this and the focus of the research question, this section is divided into three sub-sections: practical considerations, methodological considerations, and ethical considerations.

3.3.1 Practical Considerations

In 1997, Hanna *et al.* set some guidelines for usability testing with children, based on their work as usability engineers at Microsoft [18]. The proposed guidelines offer practical advice on how a researcher can set up a child-friendly testing environment and schedule usability tests in laboratories with children. These guidelines are categorized according to a four-phase test facilitation process: set-up and planning, introduction, during the test, and finishing up [18]. Many of these guidelines became important for design testing and evaluation with children. Several of the studies, particularly [5, 15], have applied the guidelines developed by Hanna and her colleagues.

Markopoulos and Bekker [7] question some of the advice or guidelines described by Hanna *et al.* [18]. Hanna *et al*. categorize children into three age groups related to their behavior during a test; preschool (2–5 years), elementary school (6–10 years) and middle-school (11–14 year). Hanna *et al*. experienced that children under 12 years of age are not able to think aloud. However, Markopoulos and Bekker experienced that

"S*ome children may not be used to speaking up to adults and may be less likely to report usability problems. Extroversion and verbalization skills are thus important variables to control.*" [7]. Moreover, [18] states that children can concentrate for about 30 min, but [7] experiences that children (9–12 years of age) can enjoy longer sessions (45 min or more). Furthermore, [7] discusses the characteristics of children that impact the usability test. These are, capacity and inclination to verbalize, capability to concentrate, children's motivation, ability to adjust to strange environments and surroundings, trustworthiness of self-report, ability for abstract and logical thinking, monitoring progress towards a goal, gender differences, knowledge of language and concepts, and knowledge and skills about computer and interactive systems' use. These characteristics of children are important practical considerations for usability tests involving children.

"*If time stood still, and technology and children never changed, the original work by Hanna, Risden et al. would no doubt still be as valid now as it was then. But as we all know, nothing stays the same, and in the dynamic area of interactive technology and children, change is inevitable and rapid.*" [6]. For example, many new technologies cannot be tested in so-called usability labs, because technologies are becoming more mobile, and interactive [6]. Today, it is more customary to conduct testing in a child's contexts and make so-called *field visits*, which, compared to lab-based tests, provide better and more useful results as the children can express themselves more freely in their natural environment [14]. Existing guidelines targeting usability testing with children need some makeover: "*The original guidelines are essentially still highly relevant because, after all, a child is still a child. However, there are three areas where it appears, in the light of the changing times, that some adjustment is needed: these are timing, screening, and participation. Also, as there is now new knowledge about interactive technology and children, there are some additional guidelines to be aware of. In the original work, the researchers held usability tests between 30 min and an hour long. In our experience, and that of many others, this now seems rather liberal. Maybe children have shorter attention spans than they used to, but modern young children can often concentrate for only very short periods—as short as 10 min—and even older children find sessions beyond 30 min problematic.*" [6]. Many such considerations are also essential when testing with adults since they also have a need for security and tendency to lose concentration during the process.

3.3.2 Methodological Considerations

Usability testing includes various methods [7, 14]. Some of the methods involving children in interaction design are the verbalization methods (including think-aloud, picture cards method, active intervention, post-task interview, robotic intervention, constructive interaction, and peer tutoring), the wizard of oz method, survey methods, diaries, and inspection methods (including Nielsen's usability heuristics, problem reporting, heuristics for evaluating fun, games and websites, persona-based evaluation, software engineering with enterprise models (SEEM method) [19]. Think aloud is a common method; however, "children below the age of 12 are likely to be unable to think aloud" [7].

Several studies have focused on examining ways that provide the best results when testing usability with children [15]. Researchers have measured, among other issues, how many problems various methods have caused. *"[The] practices vary from simply asking*

children what was successful and what not, or asking what were the 'bugs' of the software tested, to analyzing the stories children make with a storytelling program." [7]. *The think-aloud test* is the most widely used method [15]. The differences between *think-aloud* and *constructive interaction,* the two variations, is that *think-aloud* involves only one test person whereas *constructive interaction* involves two test persons who need work together to solve different problems by using an interactive system [14]. Studies showed that *constructive interaction* is the variation that identifies the most problems [14]. It is, however, important to consider the process of selecting the participants: "*Constructive interaction with pairs of children knowing each other identified more problems (on all severities) and specifically more critical problems."* [15].

Nowadays, children collaborate online for academic activities and games; the methods applied for the evaluations of interactive technology should take account of children exploring interactive systems beyond the restrictions of locations and time [6].

Children are becoming independent and gaining increasing autonomy in accessing, personalizing and owning interactive digital systems. So, there is the transformation of power, confidence, motivation and feeling of comfort. Previously, the participation of a teacher or a parent was identified as "comfortable" during tests with children [18]. About a decade later, it is claimed that such a guideline for methods does not apply anymore; children can independently explore technology [6]. Read points out that it is the location of the test and the children's age and language development that should determine which methods the researcher chooses [17].

3.3.3 Ethical Considerations

The ethical considerations in testing usability appear to be pioneered by and emphasized in the articles contributed by Read [6, 17, 20]. The articles address the following issue*: "In the IDC [Interaction Design and Children] literature very few researchers have documented how they have concerned themselves with the rights and feelings of children within the context of research using participatory design."* [20]. Furthermore, it is importance to acknowledge the children for their participation and tell them what their participation has contributed to. This is first and foremost because children, as well as anyone else, have a right to know the purpose of their involvement. Read and Markopoulos "*suggests an approach in which the research team deliberately and critically examine why they are engaging with children and then ensure that the children have as much knowledge as possible in order to decide whether or not to participate.*" [20, 21] Children should, therefore, be informed about how their ideas will be used and what happens if some of their ideas contribute to the success of the company or the investigation. Another reason for why it is important to inform children about the purpose of their participation is that "*It is generally considered that the more meaningful participation is for children the more beneficial the activity is across all aspects*" [20].

Read [17] emphasizes that it can be difficult for researchers to practice ethical considerations related to involving children in a design process. The author underlines the importance of making clear that the children understand the reason for their participation in a test, "*because if the things cannot be reasonably justified, the evaluation should not take place.*"[17]. Read [17] suggests that "*Children should be able to use*

methods they can relate to and understand so their contributions can be meaningful to them as well as to the adult evaluator."

4 Discussion and Conclusion

The overarching goal of this paper was to establish an overview of usability studies carried out targeting the inclusion of children. The questions guiding the literature review concerned the *motivation* for conducting usability tests with children and, also, the kind of *methodological*, *practical*, and *ethical* considerations that should be considered when involving children in usability studies. The information was provided through a literature search process, which was challenging in terms of finding the correct combination of keywords. Optimally, the investigation should enable a search based on different age groups, but this was found difficult, for example, when applying search phrases such as "primary school children," the hits were either irrelevant or too few. However, as keywords such as "primary schools" or "elementary school" are not universally uniform, the authors suggest that in the child interaction design discipline age groups, e.g. 3–5, 5–12, 13–18, should be included as past of keywords in scientific publications. [3] highlight the issue of considering children's developmental needs when they are included in the design process. In addition, we suggest that this should include needs related to children's cognitive as well as social, emotional, and motor abilities. This kind of considerations respect and value children's different need for support and scaffolding [3]. Furthermore, [3] state that teenager "differ enough from children developmentally that design with teenagers should be considered separately from that of children." Thus, it should be emphasized that the outcome of the current literature review has identified that the children that have been included in the different studies are grouped into *teenagers* (13–18 years of age) and children younger than teenagers (5–12 years of age) by child-computer interaction researchers [3, 22].

It is notable that not many studies from the literature review include younger children as a participant group in usability inquiries. [23] carried out a study on digital playful learning, including 55 children, 3 to 5 years-of-age in an early years educational setting. The authors investigated different methods developed for children, but not necessarily for young children. The method they applied was the Cooperative Inquiry, among others the Bags of Stuff technique [24] and the Mixing Ideas technique [25]. The results suggested that when working with younger children, it is necessary to make efforts to understand the children and their conceptual framework before engaging in design activities. Furthermore, the findings highlighted that young children need support in their creative expression.

Children have their standards and strong opinions, which means that adult assumptions are not sufficient when developing interaction designs for children. Accordingly, several of the authors of the reviewed articles consider an adaptation of the techniques traditionally used in usability testing with adults important. This includes guidelines for the settings, how to divide the children in testing situations, and the duration of the test. In the same way, many of these guidelines also apply to testing with adults, in particular regarding parameters such as concentration, language skills and comfort. Nevertheless, although

recommendations apply to, age wise, a broad range of participants, there are inherent differences. For example, children's attention span is most often shorter compared to adults' and would, therefore, require breaks more frequently than an adult participant. Furthermore, when testing with children, a special focus on the ethical considerations since, for example, children can have a hard time understanding why they are participating in the test, and they can become upset during the performance of the test.

The findings from the literature review suggest that when working with children, preparations have to be carefully considered to engage children in a usability test situation. Creating a common ground is an essential factor the communication between the participant and the researcher, and setting the right ethical frame for the sessions can strengthen the quality of the design test situations.

References

1. Zimmerman, J., Forlizzi, J., Evenson, S.: Research through design as a method for interaction design research in HCI. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 493–502. ACM, New York (2007)
2. Sharp, H., Rogers, Y., Preece, J.: Interaction Design: Beyond Human-Computer Interaction. Wiley, Chichester (2007)
3. Fails, J.A., Guha, M.L., Druin, A.: Methods and techniques for involving children in the design of new technology for children. Found. Trends® Hum. Comput. Interact. **6**(2), 85–166 (2012)
4. Read, J.C., Markopoulos, P.: Child–computer interaction. Int. J. Child-Comput. Interact. **1**(1), 2–6 (2013)
5. Druin, A.: The role of children in the design of new technology. Behav. Inf. Technol. **21**(1), 1–25 (2002)
6. Read, J.C., Markopoulos, P.: Understanding children's interactions: evaluating children's interactive products. Interactions **15**(6), 26–29 (2008)
7. Markopoulos, P., Bekker, M.: On the assessment of usability testing methods for children. Interact. Comput. **15**(2), 227–243 (2003)
8. Druin, A.: Cooperative inquiry: developing new technologies for children with children. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 592–599. ACM, New York (1999)
9. Marco, J., Cerezo, E., Baldassarri, S., Mazzone, E., Read, J.C.: Bringing tabletop technologies to kindergarten children. In: Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology, pp. 103–111. British Computer Society, Swinton (2009)
10. Veale, A.: Creative methodologies in participatory research with children. In: Greene, S., Hogan, D. (eds.) Researching Children's Experience: Approaches and Methods, pp. 253–272. SAGE, London (2005)
11. Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G.: Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann. Intern. Med. **151**(4), 264–269 (2009)
12. Glaser, B.G.: The constant comparative method of qualitative analysis. Soc. Probl. **12**, 436–445 (1965)
13. Rounding, K., Tee, K., Wu, X., Guo, C., Tse, E.: Evaluating interfaces with children. Pers. Ubiquit. Comput. **17**(8), 1663–1666 (2012)

14. Khanum, M.A., Trivedi, M.C.: Exploring verbalization and collaboration during usability evaluation with children in context. IJCSI Int. J. Comput. Sci. Issues **10**, 485–491 (2013)
15. Als, B.S., Jensen, J.J., Skov, M.B.: Comparison of think-aloud and constructive interaction in usability testing with children. In: Proceedings of the 2005 Conference on Interaction Design and Children, pp. 9–16. ACM, New York (2005)
16. Tasos, S.: Integrating Usability Engineering for Designing the Web Experience: Methodologies and Principles: Methodologies and Principles. IGI Global, Hershey (2010)
17. Read, J.: Children as participants in design and evaluation. Interactions **22**(2), 64–66 (2015)
18. Hanna, L., Risden, K., Alexander, K.: Guidelines for usability testing with children. Interactions **4**, 9–14 (1997)
19. Markopoulos, P., Read, J.C., MacFarlane, S., Hoysniemi, J.: Evaluating Children's Interactive Products: Principles and Practices for Interaction Designers. Morgan Kaufmann, San Franscisco (2008)
20. Read, J.C., Fitton, D., Horton, M.: Giving ideas an equal chance: inclusion and representation in participatory design with children. In: Proceedings of the 2014 Conference on Interaction Design and Children, pp. 105–114. ACM, New York (2014)
21. Read, J.C., Horton, M., Sim, G., Gregory, P., Fitton, D., Cassidy, B.: CHECk: a tool to inform and encourage ethical practice in participatory design with children. In: CHI 2013 Extended Abstracts on Human Factors in Computing Systems, pp. 187–192. ACM, New York (2013)
22. Iversen, O.S., Smith, R.C.: Scandinavian participatory design: dialogic curation with teenagers. In: Proceedings of the 11th International Conference on Interaction Design and Children, pp. 106–115. ACM, New York (2012)
23. Borum, N., Brooks, E.P., Brooks, A.L.: Designing with young children: lessons learned from a co-creation of a technology-enhanced playful learning environment. In: Marcus, A. (ed.) DUXU 2015. LNCS, vol. 9188, pp. 142–152. Springer, Heidelberg (2015). doi: 10.1007/978-3-319-20889-3_14
24. Fails, J.A., Druin, A., Bederson, B.B., Weeks, A., Rose, A.: A child's mobile digital library: collaboration, community, and change. In: Druin, A. (ed.) Mobile Technology for Children: Designing for Interaction and Learning. Morgan Kaufmann Publishers, Burlington (2009)
25. Guha, M. L., Druin, A., Chipman, G., Fails, J. A., Simms, S., Farber, A.: Mixing ideas: a new technique for working with children as design partners. In: Proceedings of Interaction Design and Children 2004: Building a Community, pp. 35–42 (2004)