Please note that this is the final draft post-refereeing print. For the publisher's version and for citation purposes, please use: Vanden Abeele, V., Zaman, B., De Grooff, D. (2012). User eXperience Laddering with Preschoolers: Unveiling Attributes and Benefits of Cuddly Toy Interfaces. Personal and Ubiquitous Computing, 16 (4), 451-465.

User eXperience Laddering with Preschoolers: Unveiling Attributes and Benefits of Cuddly Toy Interfaces

Vero Vanden Abeele

e-Media Lab, GROUP T, Leuven Engineering College / CUO (Association KULeuven), Vesaliusstraat 13, 3000 Leuven, Belgium

+32 16 30 10 30

Vero.Vanden.Abeele@groept.be

Bieke Zaman

CUO,IBBT/KULeuven, Parkstraat 45 bus 3605, 3000 Leuven, Belgium

+32 16 32 36 51

Bieke.Zaman@soc.kuleuven.be

Dirk De Grooff

CMC-CUO, IBBT/KULeuven, Parkstraat 45 bus 3605, 3000 Leuven, Belgium

+32 16 32 31 12

Dirk.DeGrooff@soc.kuleuven.be

Abstract. In this paper, we suggest *laddering* as a promising empirical method to evaluate the impact of tangibility on young children's user experiences. In the first part of this paper, we explain what laddering is. We explicate the conceptual foundations of laddering, discuss the typical laddering interviewing technique and focus on the laddering data treatment. Then, we argue why laddering might be especially valuable in a context of User eXperience evaluations of tangible and embedded interfaces with children. In the second part of this paper, we present a case study, comparing three cuddly toy interfaces, and we demonstrate how laddering can be used with preschoolers to explain preferences between these tangible interfaces. The case study confirms that laddering can contribute to verifying the assumed benefits of tangibility. Laddering revealed how specific cuddly toy attributes as opposed to t non-cuddly toy attributes lead to specific benefits for the young participants. However, contrary to research findings from developmental literature, only children aged five years and older proved to be capable of performing as full laddering respondents.

Keywords: *User eXperience, research methods, laddering, preschoolers, children, tangible interaction, cuddly toy interface, evaluation*

Introduction

Embodied interaction, tangible manipulation and physical mediation of digital data are the recent focus of many HCI researchers, carried out under the label of 'Tangible Embedded Interaction' (TEI). Donald Norman [1] addresses this trend as physicality or "the return to physical devices, where we control things by physical body movement, by turning, moving, and manipulating appropriate mechanical devices". TEI is considered as more natural and to be preferred over a cognition driven approach, associated with personal computing paradigms [2-4]. Tangibles are often associated e with a more affective interaction [5, 6]. TEI is also found to stimulate social interaction and sharing [4, 7]. Concerning interaction design for children, TEI is considered to enhance play and stimulate learning [8-10]. Not surprisingly, in the last years, many creative TEI-projects have been carried out to augment children's play experience and learning.

As the field expands and addresses new challenges, it is important to be clear about the main motivations and assumptions regarding the benefits of tangible interaction and to seek empirical evidence for these. Unfortunately, many research articles on TEI projects that report on the 'positive effects' of tangibility go no further than describing participants' mostly enthusiastic experiences with tangibles. Moreover, it appears that often little attention is paid to the user evaluation method. Consequently, it becomes hard to make solid claims about the results' reliability or validity with respect to the 'tangibility'. Although these user studies provide interesting, exploratory and inspiring qualitative information, the evaluation results are often equivocal; hypotheses or research questions of researchers are not properly addressed. In order to suggest design improvements regarding the user experience, one should be careful in drawing conclusions from children's enthusiastic play behaviours and reactions. Researchers should always critically question whether the enthusiasm of participating children is really caused by a positive user experience or by something else. What is often lacking in TEI research papers, is a thorough description of the user research method, how it has been tailored to a young audience, and a reflection upon the results with a special attention to the limitations of the study [11, 12]. Only by doing so, we can investigate the true impact of tangibility on children's user experience and developmental processes.

In sum, in order to support the development of the TEI research field, it is necessary that empirical studies adequately affirm the design improvements and technological innovations. User research on TEI, certainly in combination with children, lacks empirical validation on whether it is really the ‘tangible’ or ‘embedded’ interaction that is causing the positive effects. To make reliable claims about the effects of tangible or embedded interaction, adapted research designs and methods are needed. In this paper, we suggest such an adapted research design, namely laddering, as a promising empirical method to evaluate the impact of tangibility on young children’s user experiences.

In the first part of this paper, we explain what laddering is. We explicate the conceptual foundations of laddering, discuss the typical laddering interviewing technique and focus on the laddering data treatment. Then, we argue why laddering might be especially valuable in a context of User eXperience evaluations of tangible and embedded interfaces with children.

In the second part of this paper, we present a case study, and demonstrate how laddering can be used to explain preferences between cuddly toy interface prototypes with preschool aged children. A cuddly toy interface consists of a plush toy, enriched with hidden acceleration sensors, gyro sensors en pressure sensors. These sensors allow for mapping physical manipulations with the plush toy unto the movements and actions of an in-game character in a 3D world.

What is laddering?

The conceptual foundation of laddering: Means-End Theory

Laddering originates in consumer research and relies heavily on Means-End Theory as proposed by Gutman [13]. Means-End Theory states that people choose a product because it contains attributes (the means) that are instrumental to achieving the desired consequences and fulfilling values (the ends). In other words, users’ product choices and consumer behaviour are dependent on how they perceive certain product *attributes* as most likely to deliver certain desired *consequences, which* are seen as beneficial towards their individual *values*[1] . The common generic Means-End Chain, therefore, consists of attributes (A), consequences (C) and values (V).

Attributes -> Consequences -> Values

With consequences, one either points to positive consequences that a consumer seeks from a product, or negative consequences that a consumer avoids by using a product. Clearly, a consumer desires to derive benefits from product use. Therefore, the term ‘benefits’ is often used as a synonym for ‘consequences’.

Although in consumer research, a three-step (A-C-V) model is often sufficient [15], the model has been further refined by [16] as a six-level model (see figure 1), with the following steps: concrete attributes (CA) -> abstract attributes (AA) -> functional consequences (FC) -> psycho-social consequences (PSC) -> instrumental values (IV) -> terminal values (TV).

Concrete attributes are those graspable, directly perceivable aspects of a product, such as a colour, a certain material, or size. Abstract attributes are those intangible features of a product, not directly perceivable characteristics, such as the styling or the level of convenience.

Contrary to attributes which remain invariant and are not dependent on the judgment of the individual product user, consequences are directly tied towards the use and the judgment of the individual. Consequences unfold in the interaction between the user and the product. Consequences provide a blend of the product and the individual (see figure 1). Functional consequences are situated at the usage level, as exemplified by the statement “because that [product attribute] makes me go faster”. Psycho-social consequences exceed the usage level and list the consequences for social, psychological or psycho-social level, e.g., “because that [product attribute] makes me feel better about myself” or “because my friend thinks it [product attribute] is cool”.

[1] In fact, Means-End Theory closely parallels Expectancy-Value theories, as put forward by Fishbein and Ajzen, and is built upon by communication media scholars via Uses &Gratifications paradigm. Readers that would like to read more about the similarities between these models, especially with relation to ‘likeability for preschoolers’ are referred to [14].

Finally, at the value level, individuals no longer mention the product, but talk about the values, morals and norms they uphold, e.g., "because I think it is important to take care of others" or "because it is important to save the planet" which are both extrinsic or instrumental values. These values still allow serving a purpose. Terminal values are values such as "happiness", "peace-of-mind", which are intrinsically oriented.

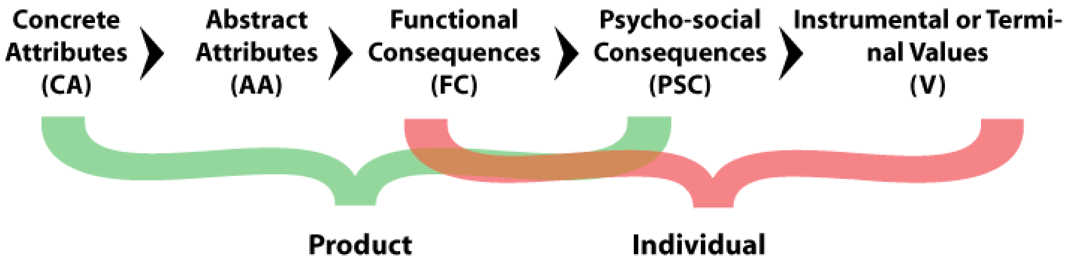

Figure 1. A six-level model for Means-End chains.

The Means-End approach foresees in a variety of methods to observe and question users about their choices for certain product attributes and desired outcomes [17]. One particular method for interviewing and the resulting data analysis is *laddering* [18]. Laddering gained popularity within consumer research as a powerful approach, allowing for segmentation of the customer audience, advertising and marketing purposes [19]. Laddering proved to be superior to other elicitation methods as it increased the explanatory power with regard to choice option attractiveness [20], i.e., Laddering is the best method for understanding why people chose one product over another

In the following paragraphs, we will explain the interviewing technique, the qualitative as well as the quantitative data treatment that constitute the laddering method.

Laddering as an interviewing technique

Laddering is most commonly known as a specific in-depth interviewing technique. The laddering interview consists of two phases. The first phase is the 'Attribute Elicitation phase'. During that phase, the interviewee is prompted to identify salient attributes of a certain product or product class as explanatory for his/her preference among several products or product classes. Different techniques are possible for attribute elicitation, such as triadic sorting, free sorting, ranking, direct/free elicitation, etc. For an excellent overview of different possible techniques we refer to [18, 21, 22]. Perhaps, the best known technique for attribute elicitation is triadic sorting, based upon the Repertory Grid Technique [23], also used within Human Computer Interaction by [24-27]. The Repertory Grid technique results in a list of attributes, ranked or sorted by the interviewee.

After this initial attribute elicitation phase, the actual 'laddering' takes place. During this phase, the interviewer will try to reveal the interviewee's product attribute related ladder by asking: "Why is this attribute important to you?". In other words, the interviewee is asked to motivate his attribute selection by explaining the related, anticipated and favoured consequences. The interviewer, however, does not rest at the consequences level and keeps prompting "Why?". This way the interviewer aims for an expression of how these consequences serve personal values.

This probing typifies a laddering interview, during which the question "Why is that important to you?" is repeated as many times as needed to reveal all possible elements of the ladders. By probing into the reasons why, the interviewee will 'climb up the ladder'. The goal of laddering -as with all Means-End approaches- is to identify and understand the linkages across the range of attributes, consequences and values, rather than deriving an inventory of the attributes, consequences or values. Through this insight in the relationships between attributes and values, researchers can understand the meaning that (product) attributes hold for users. The real interest therefore lies in the association networks; the meaningful couplings between attributes, consequences and values.

Laddering as data treatment

Although many researchers only refer to laddering as a special interviewing technique, one might easily forget that full laddering entails both an interviewing technique and a specific procedure for data analysis [15]. The laddering data analysis involves a qualitative and a quantitative phase.

In the qualitative phase, the laddering interviews are transcribed and elements (i.e., attributes, consequences, values) are coded. This phase relies on skills that are standard in qualitative research and content analysis. Reliable coding (i.e., interpreting, categorizing and labeling the data while bearing in mind the A-C-V conceptualization), is necessary to define those core elements that emerge from the interviews. Categories of elements need to be broad enough to get replications, e.g., this categories most be mentioned by more than one respondent. Yet, these categories should not be so broad that all meaning is lost.

Additionally, laddering allows for crossing over from qualitative data gathering to quantitative data treatment. Once that the core elements are defined and labelled, the individual ladders can be re-inspected and decomposed based upon these codes. All individual ladders are entered into what is called a Summary Score Matrix (SSM), summarizing the data of all the interviewees. At this point, laddering becomes a quantitative research effort. From the SSM, an Implication Matrix (IM) is constructed, which lists the amount of direct and indirect links between two elements in the ladders. A direct link occurs when two elements follow one another directly, an indirect link refers to two elements that are positioned in the same ladder but that are not positioned adjacently. The sum of direct and indirect links between the elements expresses the strength of the link between these elements.

In order to define dominant linkages, the researcher has to decide what can be considered a dominant link. The decision is made quantitatively and based upon cut-off levels. Links that fall beneath cut-off levels are ignored because they are too marginally related; those that reach that predefined level are kept. From the quantitative laddering data analysis, dominant perceptual orientations emerge that do no longer represent individual ladders but Means-End Chains (MECs) at the aggregate level. Such dominant relationships are mapped out in a Hierarchical Value Map (HVM), see figure 4 as an example. This map gives a visual overview of which attributes link to which consequences and values based on the number of direct and indirect links in the Implication Matrix. This HVM is the ultimate goal of a laddering study as it provides an overview of dominant perceptions, providing insight into the motives for using a certain product.

Why laddering for User eXperience evaluations of tangible and embedded interfaces with young children?

As was argued in the introduction of this paper, we should deliberately question and critically reflect upon the reasons why and the way in which we select our methods. In this paragraph, we will therefore explain our motivation to use laddering for User eXperience (UX) evaluations of TEI applications with very young children. Three topics form the thread throughout our motivation of laddering: user experience, TEI and young children.

Laddering in a User eXperience evaluative context

In the past years, the HCI research domain has been characterized by a number of studies, workshops and SIGs [28-34] that aimed to contribute to the dearth of frameworks, theories and methods to facilitate approaches to design products that allow for pleasurable, enjoyable, and entertaining interaction. One particularly relevant framework is it that of Worth-Centered Design, proposed by Cockton. The author and others [35-38] advocate that designing user experiences is connecting a product's means to the user's ends. As a result, these authors conceive the aesthetics of user experience as unfolding in the interaction between the product 's means on the one hand and the ultimate ends of the individual on the other. Laddering is one possible research method that can reveal these means-end chains to grasp the aesthetics of User eXperiences.

Moreover, laddering is a method that corresponds well to ISO's conceptualization of User eXperience. ISO defined User eXperience as "*a person's perceptions and responses resulting from the use or anticipated use of a product, system or service*"[39]. Approaching UX from ISO's perspective, we consider experiences as *individually constructed* and thus explicitly address the subjective nature of technological experiences, hereby also following one of the UX pioneers Hassenzahl [28]. From a methodological point of view, laddering equally aims at analysing people's experiences by relying upon individual self-reports. We will focus on one type of responses that are the *judgmental responses* which are affect-driven evaluative responses to (an interaction with) a stimuli [40]. We should hereby acknowledge that each affective response is further modified by classification and comparison processes in which "*people construct the apparent product character based on the particular combination of 'product features and their personal standards and expectations*" and that a "*personal standard most likely consists of other objects the product can be compared to*." [28]. More particularly, when judging products, people typically create bipolar (e.g., easy versus difficult) constructs when comparing the product features against the features of the product alternatives [41, 42]. This judgmental comparison process then results in "an expression of opinion about the *dominance* (importance, preference or likelihood) of one element over another". Again, laddering, -which is from origin not a HCI but a Consumer research method-, corresponds surprisingly well to our conceptualization of User eXperience. As we will explain later, laddering interviews typically start by letting the respondent (1) judge and (2) explain (3) significant differences between products or product classes (4) in terms of preferences.

Laddering in a TEI evaluative context

The second motivation for laddering refers to the methodological issues that surface when researching TEIs for children. In general, we can discern different methodological issues regarding research on children's likes and dislikes according to the research phase within the design process. In the early design phases, researchers are inclined to rely on creative methods to inform the design [43] such as cooperative inquiry [44], drawings or low-tech prototyping [45]. In the middle and late design phases, user-based evaluative methods are prominent such as the Problem Identification Pictures Cards method (PIPC) [46], (spontaneous) thinking-aloud [47], constructive interaction [48], peer-tutoring [49, 50] or the Fun Toolkit (Smileyometer, Funometer, Again-Again meter [51, 52]). As opposed to the methods of the early design phases which are mainly formative (shaping or inspiring the design), the middle and late design methods are summative (evaluating or testing the design). Whereas in the beginning of a design process, the main research aims are dealing with finding innovative and age-appropriate *design ideas for future products*, the purposes in the last stages are rather of the type of *benchmarking and evaluation*: deciding on product preferences, making comparisons between designs, assessing whether user requirements have been achieved, and detecting usability problems or a few key attributes to improve towards an optimal user experience (see also [53]).

In the case of TEI research, formative methods would focus on how the tangible interaction could be conceptualized and designed whereas evaluative methods would verify whether the tangible interaction really serves the benefits it was designed for. Summative methods can for instance be used to compare one product with a tangible interface versus the same product but with a more traditional interface (see for instance [54]).

In this paper, we will add laddering to the list of evaluative methods that can be used in the middle and late design phases when typically some test materials are available, be it sketches, models, scenarios, mock-ups or prototypes ([45, 53]). laddering will help to benchmark various TEI or non-TEI prototypes against each other. However, in the case of benchmarking and evaluating tangible interfaces, it is not sufficient to reveal which alternative scores best. Laddering not only aims at revealing which prototype was preferred but it also allows for understanding into the reasons *why*. Consequently, laddering points to those concrete product attributes that are related to (or may be improved in order to serve) positive values. For TEI research, laddering thus answers the question of to what extent the tangible or embedded attributes deliver intended consequences and value to the overall user experience and why.

Laddering in a context of evaluations with children

Finally, we should question whether laddering is an appropriate research method for young children. In this paper, we deal with the preoperational age group of preschoolers as labelled by Jean Piaget, encompassing children from two to seven years old. These children cannot yet read or write and have not yet started infant school. Nevertheless, according to the literature, children from the preoperational stage already show cognitive and behavioural capabilities needed to be adequate respondents in a laddering interview.

Language

One of the most prominent capabilities that children from the preoperational stage acquire is *language* [55]. Preschoolers' language has reached a sufficient level for the adequate use of words to describe actions or represent objects [56, 57]. The same goes for their *pragmatics*, or "children's practical ability to use language to communicate with others in a variety of social contexts" [56] and preschoolers' ability to deal with the complexities of open-ended questions [58]. As for the *ability to talk about experiences* in general and media or technology experiences in particular, preschoolers have acquired the skills to think and talk beyond the here and now. Yet, in general, young children's cognitive abilities are maximized in situations where children are given tasks that are simplified and relevant to their everyday lives [56]. In the case of laddering, we should therefore ensure that the technologies evaluated are meaningful objects used in a familiar context-of-use.

Awareness of psychological causes

Apart from their language and pragmatics, young children's *awareness of mental states and their understanding of psychological causes* are other relevant cognitive capabilities for laddering interviews in which the interviewer probes for the underlying reasons of positive or negative user experiences. When children begin to reach the age of three, their ability to speak about mental states grows [59, 60] as well as their ability to think of events and behaviours in terms of causal relations [61] . Miller & Aloise [62] came to the same conclusions for two- to five-year-olds who proved to be able to appropriately and spontaneously use words referring to their internal state while acknowledging for the psychological causes of behaviour and events. The mental states children mentioned, seemed to involve past, present or future, and refer to 'cold' aspects of mind such as thought, memory, attention and perception or 'hot' aspects such as emotions, motivation and personality [62]. Miller & Aloise concluded that these young children's expressions of mental states were "not limited to their own states or the immediate, concrete situation." More particularly, young children can talk about mental states that they are not actually experiencing or mental states of others [57, 63].

Based on these theoretical insights, we may assume that laddering should be possible with respondents aged two to seven years old. Nevertheless, this assumption is taken with care and will be critically evaluated in our case study as the preoperational age group already provides a vast age span. Although theoretically these children share many similar capabilities, there is more variation in the preoperational stage than may spring from Piaget's stage theory [64]. In other words, although we often refer to the 'typical' achieved intellectual and language developmental capabilities of the preoperational children, one may not forget the individual and cultural aspects cause in-group differences.

Evaluating the UX of cuddly toy Interfaces with preschoolers: a case study

In the following part of this paper, the process of laddering is discussed via a specific case of the UX evaluation of three cuddly toy interfaces. By this case study, we aim for investigating *the feasibility and usefulness of laddering when performed with young children in order to evaluate tangible interaction.* If laddering proves to be feasible with this young target group, we should get an insight into overall product preferences, and understand how product related characteristics can be improved to enhance the user experience by meeting children's needs or values. More specifically, in this case study, laddering was performed in order to reveal children's preferred prototype, as this would be the prototype the developers were to design further. Furthermore, in order to inform the design of the preferred prototype, laddering also had to provide a better understanding of which TEI movements and mappings would be preferred by the children and which other non-tangible attributes would be mentioned by the preschoolers as important criteria for deriving benefits.

Positioning the TOEWIE case - A Cuddly Toy Interface

TOEWIE is an on-going research project to design a 'cuddly toy interface' for preschoolers to navigate and interact in a 3D game environment. Our cuddly toy interface bears resemblance to the sympathetic interface developed by Johnson et al. [65] or the SenToy developed by Paiva et al [66]. Similar to these research projects, physical manipulations with a plush toy or puppet are mapped unto the movements and actions of an in-game character in a 3D world. Nevertheless, these studies seem to limit their focus to usability issues (rather than also studying UX related issues) or they relied upon methods that are only appropriate for older children. With our case study, we aim at relying upon a preschooler-friendly methodological approach that allows for studying both usability and User eXperience related issues.

Our first aspiration was to make our cuddly toy interface more useable and more likeable as compared to standard modes of navigating and interacting in a 3D world. We hereby explicitly chose for a comparable research design in order to verify the assumed added value of the cuddle toy interface. In [67, 68] we conducted an experimental study of the likeability and usability of TOEWIE versus a keyboard interface. Contrary to our hypotheses, the results of this study indicated that the cuddly toy interface was neither more likeable nor more usable than the keyboard interface. Especially the necessary physical hopping movement appeared as less liked by preschoolers. However, the preference for the keyboard could also have been explained from a psychosocial perspective; preschoolers wanted to interact with computers as their older siblings did. The results of this first study raised concerns with regards to which attributes of the cuddle toy interface prototypes were holding which benefits for preschoolers, and to what extent the tangible interaction was efficacious?

In order to address these research questions, the team designed three different cuddly toy prototypes that each incorporated a different cuddly toy interface with a different mode of physical manipulation, with an associated game character and game world. The three different prototypes are a penguin with a snow landscape, a bird within a jungle environment and a kangaroo within the woods (see figure 2). In every game, treasures are to be collected, respectively fish, apples or leaves. In order to make the penguin walk, the child needs to wiggle the penguin (tilt from left to right). In order to make the bird fly, the child needs to move the wings up and down. In order to make the kangaroo jump, the child needs to make a hopping movement. Furthermore, when the child uses the prototype in a different, non-functional way (e.g., when making a bizarre movement), the penguin topples, the bird makes a tumble and the kangaroo makes a split. This has no further function as it does not help in collecting the treasures.

We explicitly chose for a comparable research design in which the penguin, bird and kangaroo cuddly toy interactions can be compared to each other. This allowed for investigating the added value of TEI attributes (i.e., the physical manipulations and the cuddly toy itself) and whether the tangible interaction versus other in-game attributes really accounted for the overall user experience.

Figure 2. An overview of the three different cuddly interfaces and the game worlds.

Research aims

To summarize, with this case study we aim to address the following research questions:

RQ1: Is laddering feasible with young child respondents aged two to seven years old?

RQ2: Does laddering reveal overall preferences between prototypes?

RQ3: Does laddering help to explain product preferences in terms of TEI-related attributes?

As for the first question, literature suggests that laddering is possible with young child users, aged two to seven years old. Therefore, we will investigate whether our young child participants would be able to produce ladders. Additionally, we will inspect if age mediates in the number of meaningful ladders and the number of elements within one ladder?

As for the second research question, we will verify whether laddering can start with a ranking exercise in which children are asked to reveal their preferred choice.

Finally, with the third research question, we will analyse whether laddering can help us reveal the exact benefits of tangible and non-tangible attributes of the three prototypes with respect to the user experience of preschoolers. More particularly, we will question whether laddering can help us answer the following questions. Which Cuddly Toy Interface would preschoolers prefer and which means-end chains would be decisive in their choice? Which attributes would fulfil which consequences and lead to which values? We especially aim to investigate to what extent the ‘tangible’ attributes would be decisive in children’s preferences.

Method

Participants & location

All interviews took place at an elementary school, where during holidays day camps are organized for kids of the community. The testing took place on three consecutive days. 46 preschoolers that were visiting the day camps participated in the laddering study (M= 57,54 months, SD = 13, 54 months). The youngest participant was 33 months old (almost three years old); the oldest was 86 months old (seven years old). One facilitator was present to guide the child and conduct the interview. In the same room were also one data logger and one wizard (who mastered the technology and steered the interaction behind the scenes[2], see [69, 70] for examples), they did not intervene during the interview.

Procedure

In order to provide the child with a meaningful context for the product experience evaluation, the child was invited to play with either one of the three cuddly toy interfaces at the beginning of the test. Immediately after playing a game, the child was invited to play that same game a second time, in order to assure that the child fully understood the game. The order in which each child played twice with either the kangaroo, bird or penguin games was counterbalanced to avoid order effects.

After playing with the three different prototypes, the child was invited to sit in front of screenshots of the three different game prototypes as well as the cuddle toy interfaces (see Figure 3). Again, this was done to stimulate children’s memory by providing a visible memory aid.

---

[2] None of the children realized that a human Wizard steered their interaction, a pilot study revealed that even adults did not detect this.

The laddering interview started by eliciting children's preferred choice. The 'ranking' questions were inspired by the This-or-That method [71, 68] which entails five questions that are tailored to preschoolers' language and cognitive capabilities to infer which prototype was preferred most. Each question began with *"Show me which game..."* and was then competed with*"... was most fun?" "...you would like to receive as a gift?", "...you found a little bit stupid?", "... you would like to play again?", "...you would like to take home".* At the end, a control question was added to this list and prompted the child once more to identify his or her preferred choice [71] upon the sentence: *"I'm sorry, I have forgotten, can you show me once more which game you found most fun?".*

Upon the identification of the preferred prototype, the facilitator started the laddering interview by probing "*Why?*". Contrary to the typical laddering approach as it is performed with adults, we kept the laddering interviews short. If a child remained silent, repeated itself, or indicated in any other way that it was impossible for him/her to answer, we finished the interview. On average, the interview lasted two or three minutes.

Figure 3. An overview of the interview setting, with the three different conditions positioned in front of the child.

Qualitative data analysis

All interviews were recorded and transcribed. In the next step, the transcriptions of the interviews were analyzed by both authors independently, to establish the core elements (attributes, consequences and values) from the interviews. These two lists of core elements were compared to each other, and refined into one set of core elements that both researchers agreed upon (see table 1).

Table 1. A list of the core elements.

Prototypes	
1	bird
2	penguin
3	kangaroo
Concrete Attributes	
4	phys. Flapping
5	phys. Wiggling
6	phys. Hopping
7	flying
8	walking
9	jumping
10	aest. Tumble
11	aest. Topple
12	aest. Split
13	apples
14	fish
15	graphical details
16	cuddly toy parts

Abstract attributes	
17	soft/cute/cuddly
18	healthy/tasty
Functional consequences	
19	experiencing extraordinary
20	being easier/faster
21	novel/different
22	similar/recognition
23	beautiful/cool/nice to watch
Psycho-social consequences	
24	immersion/fantasy
25	funny/humour/laughing
26	social interaction
27	mastery
Terminal Value (Conceptual)	
28	Fun

In total, we discern 12 concrete attributes, two abstract attributes, five functional consequences and four psycho-social consequences.

Prototypes 1-2-3 refer to the answer on the ranking questions and indicate the prototype that was preferred most. Attributes 4-5-6 refer to the physical manipulation, often demonstrated by the child rather than described (e.g., the child demonstrated the physical manipulation). Attributes 7-8-9 refer to the navigation mode *in the game*, but do not refer to the physical manipulation (e.g., “I preferred the bird, because it could fly”). Attributes 10-11-12 refer to the non-functional aesthetics in the game, the tumble, the topple and the split. Attributes 13-14-15 refer to aspects of the digital game environment. Attribute 16 refers to specific parts of the cuddly toy. Attributes 17-18 are more abstract attributes that children mentioned in their explanation of prototype preference.

Further, five functional consequences were derived, namely ‘experiencing extraordinary things’, (e.g., “I can sit in a tree”), being faster/easier (e.g., “That was faster”), novel/different (e.g., “I don’t have that at home”), similarity/recognition (e.g., “It is like the bird of my neighbour”) and beautiful/nice to watch (e.g., “I like to look at it”).

Finally, we were also able to detect dominant psycho-social consequences from the interviews. Children pointed towards immersion (e.g., “then I feel like a bird"), humour (e.g., "it makes me laugh"), social interaction (e.g., "I can tell it to my mommy") and mastery (e.g., "then I can do it better").

None of the ladders attained the value level. However, due to the nature of the ranking question that formed the basis of the laddering interview (*"Which did you find most fun?"),* we argue that conceptually every ladder links to the terminal value of 'fun'.

Next, the interviews were re-analyzed and coded by both authors independently on the basis of this new list of core elements. On the basis of this second resulting dataset, interreliability between coders was assessed. Cohen's Kappa was found to be at 0.77 ($p < .0.001$), demonstrating an acceptable level of agreement between coders. Table 2 shows a typifying example of a transcribed laddering interview that was administered with a boy aged 57 months (almost 5 years).

Table 2. Example of an excerpt of a laddering interview with a boy, almost 5 years old

LADDERING QUESTIONS AND ANSWERS
"I'm sorry, I have forgotten, can you show me once more which game you found most fun?".
[child points towards the bird game] *(prototype preference Bird)*
What did you like about that game?
That it can fly *(Attribute Flying)*
Why do you like flying?
Because flying, it is actually, it is very … Nice because … actually an airplane can also do a somersault
The bird flies, what do you like about flying?
Because it can do a somersault *(Attribute Tumble)*
What did you like more about it?
Because it's like I'm flying *(Consequence Immersion)*

Quantitative data analysis.

To arrive at the dominant means-end chains, we first coded each ladder on the basis of the above list of core elements. These ladders were entered into Ladderux [72] [3], an on line tool that aids the quantitative analysis of laddering. The coding of the ladders was summarized in a *score matrix*. This score matrix lists up all ladders for all individual respondents. In total, the score matrix consisted out of 79 ladders for 46 preschoolers.

---

[3] http://www.ladderux.org/lux2.1/

Secondly, from this score matrix, the *implication matrix* was drawn. The implication matrix charts the links between all elements, both direct and indirect. Contrary to the score matrix which presents data at individual level, the implication matrix represents data at the aggregate level. The laddering resulted in a total number of 223 elements, yielding 256 links (144 direct links, 113 indirect links). To determine an appropriate cut-off level (thus link strength between conditions concrete attributes, abstract attributes, functional consequences and psycho-social consequences), we based ourselves on the heuristics provided by [73, 74] suggesting a cut off level that leads to the most informative and interpretable solution, yet retains the most information. Mostly, this cut-off level lies between 3 to 5, and retains roughly two thirds of all links. Consequently, we chose a cut off level of three for all elements, retaining 162 links of the 256 links, corresponding to two thirds of all links.

As the final step in the quantitative data analysis, we drew a *Hierarchical Value Map (HVM)*, as can be seen in figure 4. The HVM visually summarizes the dominant links, and hereby reveals how the attributes of the different prototypes provide meaningful consequences for our preschoolers. The thicker the lines between two elements in the Means-End Chain, the higher the link strength, indicating that it was mentioned by a relatively high proportion of the participating preschoolers.

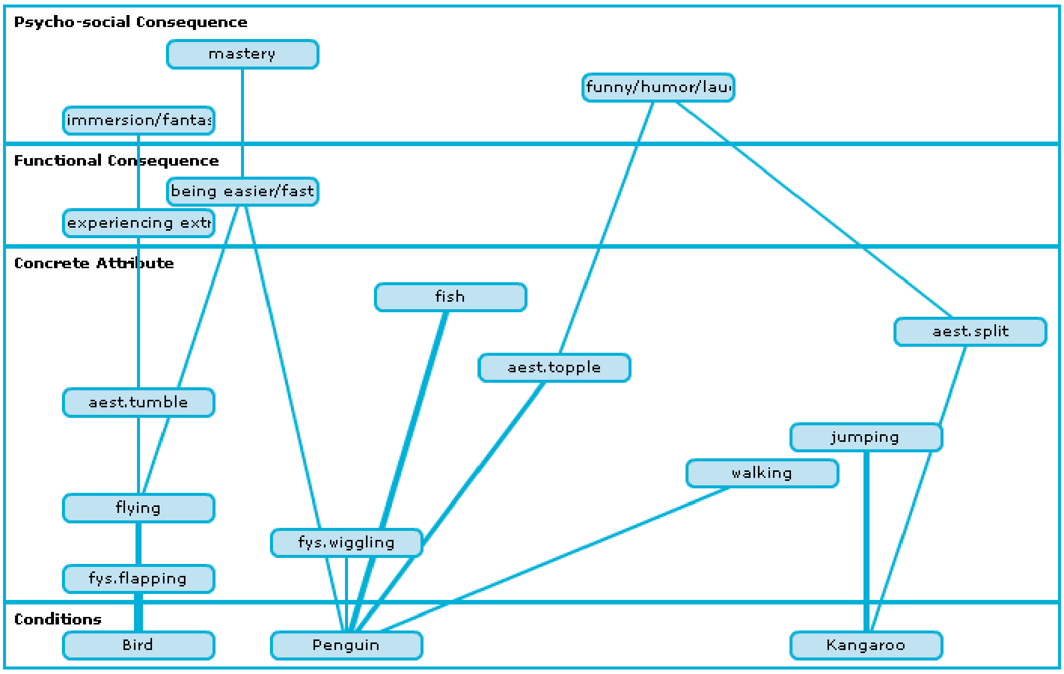

Figure 4. The Hierarchical Value Map (HVM) for the three cuddly toy interfaces.

Considering the HVM from bottom to top, the preference for each condition is linked with specific attributes which in their turn each provide meaningful consequences to the preschoolers. These consequences may be functional or psycho-social.

Results

RQ1: Is laddering a feasible method with young children aged two to seven years old?

Our first research question dealt with the feasibility of laddering with preschoolers, aged two to seven years old. We wondered whether our young participants would be able to produce ladders and questioned how age would mediate the number of meaningful ladders and the number of elements within one ladder.

The results of our case study show that in total, 79 ladders and 223 elements were derived from 46 preschoolers. This is an average of 1.72 ladders per preschooler. However, contrary to the developmental literature findings, nine children could not successfully participate in the laddering interview, and failed to produce a ladder that went beyond the choice of the preferred prototype. The mean age of these nine preschoolers was 43.67 months (SD=9.92) – about three and a half year-, as compared to the mean age of the preschoolers producing ladders, M= 60.89 months (SD = 12.17) – five year -. Further inspection revealed a strong correlation between age and the numbers of ladders that were produced by the preschoolers, r .607 ($p< 0.01$), and a strong correlation between age and number of elements, r. 469 ($p<0.01$).
Furthermore, as was expected from literature, our results confirm that children are not likely to talk about values. In our case study, none of children’s ladders attained the value level.

RQ2: Can laddering reveal overall preferences between several prototypes?

A crucial condition for laddering is the attribute-elicitation phase during which the respondent is mentioning those attributes that explain product preferences. Therefore, the child respondents should be able to express overall product preferences first. Our results show that all preschoolers, even the youngest ones, were able to indicate a preferred choice, and therefore produced a ladder that started from the preferred prototype and linked to the theoretical value of 'fun'. However, preferences were often expressed in a non-verbal way, by a child pointing to his or her preferred game.

Upon inspection of the answer to the first question "Which did you find most fun?" we noticed that the penguin is preferred by 21 children, the bird by 19 children and the kangaroo by 6 children. Although none of the three games significantly stood out in terms of first preference at the group level, the bird and the penguin game were competing for the first choice; the kangaroo game was mentioned the least by the children.

When we split the data according to gender, we see that among girls (N_{total} = 19) the penguin is preferred most (N=13 or almost 70%), next the kangaroo (N=5) and last the bird (N=1). Among boys (N_{total}=27) the bird is preferred most (N=14 or 52%), next the penguin (N=8) and least the kangaroo (N=5). Thus, although we did not find any significant statistical difference in UX preference at the group level, there is a clear trend that girls prefer the penguin and boys prefer the bird.

RQ3: To what extent does laddering help to explain product preferences in terms of TEI-related attributes?

With this third research question, we analysed whether laddering can help us reveal the exact benefits of tangible and non-tangible attributes of the three prototypes with respect to the user experience of preschoolers. More particularly, we questioned whether laddering would help us reveal what would be decisive in children's preference, which attributes would fulfil which consequences and lead to which values, and the extent to which the 'tangible' attributes would be decisive in children's preferences.

In order to answer our questions, we should first scrutinize the results of our *qualitative analysis*. From the list of core elements drawn from our laddering transcripts (see table 1), we see that from the set of 12 concrete attributes, only three referred specifically to the physical manipulation of the cuddly toy (4-5-6) and only one to the cuddly toy itself (16). Therefore, only four out of the twelve attributes could be linked directly to the cuddle toy interaction.

In order to further interpret the weight that these concrete characteristics have, we should consider the dominant ladders as they were depicted in the Hierarchical Value Map (figure 4). From this Hierarchical Value Map, the first thing to note is that the bird is the only cuddle toy interface that links to the psycho-social consequence of immersion via the dominant Means-End Chain (MEC) 'Bird' – 'phys. flapping' – 'flying' – 'aest. tumble' – 'experiencing extraordinary things' – 'immersion/fantasy'. Our results hereby suggest that children that preferred the bird seemed to value the psycho-social consequence of 'immersion'. For our preschoolers, the cuddle toy interaction of the bird was thus appreciated since it allowed for more immersion and fantasy, something which would be harder to realize without the cuddle toy interface. From the MEC, we understand that the physical flapping contributes to the attainment of the psycho-social value of immersion. From our ranking exercise, we also understand that this MEC has gendered components as mainly boys preferred the bird prototype.

Both the bird and the penguin linked to the psycho-social consequence of 'mastery' 'via the MEC 'Bird' – 'phys. flapping' – 'flying' – 'being easier/faster'- 'mastery' and the MEC 'Penguin'– 'being easier/faster'-'mastery'. These results show that both the bird and penguin interfaces allowed the child to play the game well and gave a sense of mastery. Yet for the bird this was related to the physical flapping of the wings, whereas with the penguin prototype this was not hinging upon physical interaction. The kangaroo was not linked to this consequence.
The HVM also shows that the kangaroo and the penguin offered the psycho-social consequence of funniness/humour/laughing via the MEC 'kangaroo' – 'aest.split' – 'funny/humour/laugh' and the MEC 'penguin' – 'aest.topple' – 'funny/humour/laughing'. Interestingly, the aesthetic tumble of the bird did not

contribute to humour but rather to immersion. The kangaroo only attained one psycho-social consequence whereas the other two prototypes attained two psycho-social consequences, this might explain the lack of popularity for the kangaroo.

From the inspection of the dominant MECs, we conclude that for the penguin and the kangaroo prototypes, preference is not based on physical manipulation but rather upon in-game attributes, leading to fun and mastery. For the bird prototype, the physical manipulation ('flapping the wings') contributes to a sense of mastery and to immersion in the game.

In sum, as for our three research questions, we can conclude that *laddering is possible with children aged five years and older (RQ1).* All children, also the youngest ones, proved *able to express their prototype preference (RQ2).* In order to inform the evaluation and design process of the TOEWIE case, *the laddering data provided us with insights into the reasons for preferences, allowing us to understand the impact of the tangibility on the overall User eXperience (RQ3).* From the qualitative analysis, it shows that tangibility is not the sole reason for preferring one of the three prototypes. Many other attributes are taken into consideration by the child, such as the graphical environment of the game and the non-functional aesthetics. Furthermore, from the HVM, we infer that cuddly toy and non-cuddly toy attributes work together, e.g., both the aesthetic tumble as well as the flapping with the wings contributed to immersion in case of the bird. Furthermore, not all physical movements are liked the same. Hopping and wiggling the cuddly toy did not contribute to 'immersion', nor to 'being easier/faster' whereas flapping with the wings did. Finally, we noted a strong gendered influence, which is however beyond the scope of this research paper.

Guidelines for laddering for TEI evaluations with young children

In the following paragraphs, we will reflect upon how laddering was feasible and useful in our case study to evaluate products with young children. We will discuss the adaptations we made to the 'traditional' way of laddering to make it more child-friendly or realistic in a user-centered design context. We end with critical remarks that highlight the issues for further work.

Provide alternatives

This first preliminary case study of laddering with young children showed the potential of the method. Laddering is indeed based on the premise of finding reasons for product (class) differences e.g., in terms of preferences. By providing alternatives, children can compare products or product classes, and more easily discuss their opinions and experiences.

Accept Attributes-Consequences Ladders

Our case study results show that laddering can be possible with children aged five years and older. We hereby make an important remark that we did not strive to probe until values. First of all, children are more likely to explain their preference in terms of external causes, related to the characteristics of the object entity (e.g., “This is a challenging game with difficult obstacles.”) than internal reasons, related to the characteristics of the person who made the choice (e.g., “I like to overcome challenging situations”) [62]. This implicates that in a laddering interview, children will be more likely to express attributes or product related consequences than values. Secondly, discussing (abstract) values also demands a cognitive reasoning that young children do not yet fully master. Last but not least, as we were evaluating prototypes, we also acknowledge the possible lack of values that are based upon real life experiences [75]. HCI projects often deal with new prototypes, artefacts, systems or services. In this case laddering is used to gather requirements on attributes of a product that is not yet adopted at a wide scale into society. Consequently, the Means-End Chains (MECs) that are revealed in these (HCI) contexts differ from MECs that consumer researchers find with respect to established products. In such an HCI product design and evaluation context, the laddering method should foresee that for certain consequences and values, people rely on hypotheses instead of actual user experiences within a real life context. Therefore, values listed by users might not be reliable or simply non-existent.

Avoid hard laddering

Hard laddering [76], involving putting hard pressure on the respondent to reveal consequences and values, and not stop the interview until all levels of the means-end chain are addressed. This is not without risks when the links between attributes, consequences and values might be weak because of the possible absent or restricted usage situations. The user might fabricate associations that are not the result of introspection but rather to please the interviewer. Therefore, we argue that in the case of laddering with preschoolers and/or prototype evaluations, hard laddering is likely to decrease the validity of the results of laddering, rather than augmenting them. We suggest that more weight should be given to the salient elements, rather than exhausting participants to paint a complete picture. Reaching the level of functional or psycho-social consequences already provides the UX evaluator with information about meaningful associations between the individual and the product attributes [75]. We state that it might not always be possible nor necessary to distinguish between psycho-social consequences, instrumental values or terminal values as all three indicate how attributes relate to the benefits for the individual. Finally, whereas we advice not to stress on attaining the value level within laddering, we suggest to stress more on refining the attribute level. More particularly, it is important to distinguish at the attribute level between concrete and abstract attributes [75]. Design improvements towards meaningful value or UX can only be realized if information is known about concrete attributes; abstract values are not sufficient.

Introduce a play experience as an icebreaker and memory aid

Another modification to the 'traditional' way of laddering to make the technique more child-friendly and feasible for User-Centered Design, deals with the introductory play experience. Product use should be associated to the 'real' context in which it is used. *"Laddering works best when respondents are providing associations while thinking of a realistic occasion in which they would use the product. "[...] Attention to the context of consumer behaviour provides a more meaningful context for laddering to proceed. People do not use or consume products in general; they do so in particular contexts [77]*". For prototypes, it is not always possible to perform the laddering interview in a meaningful or existing context. Nevertheless, at least the prototype experience should be recent. A preceding experience with a prototype can serve as a meaningful context to the laddering interview. This ensures that the product experience is fresh in memory, and helps the child to overcome possible cognitive limitations.

An introductory product experience may also serve as an ice breaker and provide a common topic between researcher and child to refer to. Moreover, creating a meaningful context for the laddering interview goes further than the exploratory play experience alone. It also refers to the object of experiences that stays in front of the child during the laddering interview [78]. Again, the presence of the object of experience may stimulate children's memory. If children find it hard to express their opinions in words, then they should definitely be encouraged to *show* what they are willing to say.

Keep a laddering interview short

Reynolds and Gutman recommend that each laddering study contains about 20 respondents, to arrive at over 125 ladders and at least 500 elements, assuming that an average respondent produces five ladders and that an average ladder consists of five elements. It is evident that when child respondents are involved, the interview cannot last as long and that the number of ladders and elements that Reynolds et al. propose cannot be reached. Therefore, we propose that laddering interviews in a UX evaluative context with children should remain short and not last no longer than five to ten minutes [75]. This suggestion cannot only be motivated from a child-respondent perspective but also with a user-centered design argument. Although our timing represents only a fraction of the time necessary for traditional laddering, for UX researchers, however, laddering might be only one phase of a complete design process. Consequently, full laddering with children as employed within consumer research is not realistic. Especially user-centered design, with its focus on iterative user testing with prototypes, requires an adapted, less time consuming laddering alternative.

Discussion

From a theoretical perspective, it seemed that laddering would be possible with children as young as two years old. However, in our case study we found that the youngest children were less likely to produce ladders and more likely to generate less elements within a ladder. If one aims for five elements per preschooler, linear regression analysis predicts an age of 59,57 months or five years (r. 469, b_o=-1.314, b_1=.106, p<0.01). Consequently, we suggest that laddering is only feasible with 'older preschoolers', preferably five years or older.

According to these figures, with a set of 45 preschoolers and five elements, the research study provides 225 elements, which is close to the 223 elements we found. According to Reynold and Olson, a traditional laddering study should aim for approximately 500 elements [22]. This number includes value elements as well, something we cannot expect our preschoolers to express. We are nevertheless convinced that laddering should not aim for the number of elements put forward in consumer research. We deal with prototypes and do not aim for a complete listing of all consequences and values but rather the most salient ones. However, how many elements a laddering study should aim for to claim reliable and valid results deserves further investigation and thus outlines the future efforts needed to achieve progress in this area.

Conclusion

With the emergence of new tangible interfaces, the HCI community is fully aware of the need for theories and methods for UX research and evaluation. This is especially true when one aims at revealing and understanding the benefits of tangibility. In this paper, we introduced laddering into the field of Child Computer Interaction in order to perform UX evaluations with preschoolers. This paper questioned the value and feasibility of laddering for UX evaluations of tangible interfaces with preschoolers both from a theoretical as well as a practical perspective. By our case study, we showed how laddering is instructive in understanding preferences and more particularly how it reveals the way specific cuddly toy interface aspects and non-cuddly toy interface attributes can lead to specific benefits. In other words, laddering hereby helps to critically question the assumed benefits of tangibility. As for the feasibility of laddering, the method proves to be feasible with the older children, preferably aged five years and older. We work towards more laddering projects and welcome other researchers to redo a similar case study to confirm our findings, hereby also critically reflecting about the assumptions we made. We would advice that in future work, more attention is paid to children’s background and gender (e.g., previous experience, favourite cuddle toys, ...) or the interview context and the way they may impact the results. We also suggest that future studies focus on the relation between adult and child, questioning the way in which the researcher’s characteristics may influence the results e.g., via the child’s extroversion.

Acknowledgments

We like to thank all preschoolers who participated in the laddering study. Furthermore, we like to thank the animators of the summer day camp Kattestroof and the community of Bertem, JITSOC for granting us the permission to conduct this research on their play grounds. We like to thank master students Leonard Van Domselaer (for creating the penguin) and especially Michiel Van Minnebrugge (for his wizardly patience). We also like to thank the under or above the radar work of the TOEWIE team, namely technical artists Jelle Husson and Jeroen Wauters for kakatoewie and kangaroo, Maarten Vanoverveldt and Luc Geurts for the future signal processing, and Frederik Windey for the aid in the game development… and finally Stef Desmet for his meticulous reading.

References

1. Norman, D.: The next UI breakthrough, part 2: physicality. interactions. 14, 46-47 (2007).
2. Fitzmaurice, G.W., Ishii, H., Buxton, W.A.S.: Bricks: laying the foundations for graspable user interfaces. Proceedings of the SIGCHI conference on Human factors in computing systems. pp. 442-449ACM Press/Addison-Wesley Publishing Co., Denver, Colorado, United States (1995).
3. Ullmer, B., Ishii, H.: Emerging Frameworks for Tangible User Interfaces. In: Caroll, B.W. (ed.) Human-computer interaction in the new millennium. pp. 579-601Addison-Wesley, MA (2001).
4. Dourish, P.: Where the Action is: The Foundations of Embodied Interaction. MIT Press, Cambridge, MA, US (2001).
5. Djajadiningrat, T., Wensveen, S., Frens, J., Overbeeke, K.: Tangible products: Redressing the Balance between Appearance and Action. Personal Ubiquitous Comput. 8, 294-309 (2004).
6. Wensveen, S.A.G., Djajadiningrat, J.P., Overbeeke, C.J.: Interaction frogger: a design framework to couple action and function through feedback and feedforward. Presented at the , Cambridge, MA, USA (2004).
7. Hornecker, E., Buur, J.: Getting a grip on tangible interaction: a framework on physical space and social interaction. Proceedings of the SIGCHI conference on Human Factors in computing systems. pp. 437-446ACM, Montréal, Québec, Canada (2006).
8. Antle, A.N.: The CTI framework: informing the design of tangible systems for children. Proceedings of the 1st international conference on Tangible and embedded interaction. pp. 195-202ACM, Baton Rouge, Louisiana (2007).
9. Price, S., Rogers, Y., Scaife, M., Stanton, D., Neale, H.: Using 'tangibles' to promote novel forms of playful learning. Interacting with Computers. 15, 169-185 (2003).
10. Marshall, P., Price, S., Rogers, Y.: Conceptualising tangibles to support learning. Presented at the , Preston, England (2003).
11. Xie, L., Antle, A.N., Motamedi, N.: Are tangibles more fun?: comparing children's enjoyment and engagement using physical, graphical and tangible user interfaces. Presented at the , Bonn, Germany (2008).
12. Marshall, P.: Do tangible interfaces enhance learning? Proceedings of the 1st

international conference on Tangible and embedded interaction. pp. 163-170ACM, Baton Rouge, Louisiana (2007).
13. Gutman, J.: A means-end chain model based on consumer categorization processes. Journal of Marketing. 46, 60-72 (1982).
14. Vanden Abeele, V., Zaman, B.: The Extended Likeability Framework: A Theoretical Framework for and a Practical Case of Designing Likeable Media Applications for Preschoolers. Advances in Human-Computer Interaction. 2008, 11 (2008).
15. Reynolds, T., Gutman, J.: Laddering theory, method, analysis, and interpretation. In: Reynolds, T. and Olsen, J. (eds.) Understanding consumer decision making : the means-end approach to marketing and advertising strategy. pp. 25-52Lawrence Erlbaum, Mahwah N.J. ;;London (2001).
16. Olson, J.C., Reynolds, T.J.: Understanding consumers' cognitive structures: Implications for advertizing strategy. Advertising and Consumer Psychology. pp. 77-90Lexington Books, Lexington, MA (1983).
17. Grunert, K.G., Beckmann, S.C., Sorensen, E.: Means-end chains and laddering: an inventory of problems and an agenda for research. Understanding Consumer Decision-Making: The Means-end Approach to Marketing and Advertising Strategy, pp. pp.63-90.Lawrence Earlbaum, Mahwah, NJ (1996).
18. Reynolds, T., Gutman, J.: Laddering theory, method, analysis, and interpretation, (1988).
19. Bech-Larsen, T.: Model-based development and testing of advertising messages â€“ A comparative study of two campaign proposals based on the Meccas model and a conventional approach. University of Aarhus, Aarhus School of Business, The MAPP Centre (2000).
20. Grunert, K.G., Bech-Larsen, T.: Explaining choice option attractiveness by beliefs elicited by the laddering method. Journal of Economic Psychology. 26, 223-241 (2005).
21. Bech-Larsen, T., Nielsen, N.A.: A comparison of five elicitation techniques for elicitation of attributes of low involvement products. Journal of Economic Psychology. 20, 315-341 (1999).
22. Reynolds, T.J., Dethloff, C., Westberg, S.J.: Advancements in Laddering. Understanding Consumer Decision-Making: The Means-end Approach to Marketing and Advertising Strategy, eds Reynolds, T.J., Olson J.,C. pp. 91-118Lawrence Elbaum Associates, New Jersey (2001).
23. Kelly, G.A.: A Theory of Personality: The Psychology of Personal Constructs. W.W. Norton & Co. (1963).
24. Hassenzahl, M., Trautmann, T.: Analysis of web sites with the repertory grid technique. CHI '01 extended abstracts on Human factors in computing systems. pp. 167-168ACM, Seattle, Washington (2001).
25. Karapanos, E., Martens, J.: The quantitative side of the Repertory Grid Technique: some concerns. proceedings of the workshop Now Let's Do It in Practice: User Experience Evaluation Methods in Product Development. , Florence, Italy (2008).
26. Hawley, M.: The Repertory Grid: Eliciting User Experience Comparisons in the Customer's Voice :: UXmatters, zotero://attachment/613/.
27. Fallman, D.: Integrating User Experience into the Design Process with the Repertory Grid Technique: Some Preliminary Notes. Presented at the (2004).
28. Hassenzahl, M.: The thing and I: understanding the relationship between user

and product. In: Blythe, M., Overbeeke, K., Monk, A., and Wright, P. (eds.) Funology: from usability to enjoyment. pp. 31-41Kluwer Academic Publishers, Norwell, MA, USA (2004).
29. Obrist, M., Roto, V., Väänänen-Vainio-Mattila, K.: User experience evaluation: do you know which method to use? Proceedings of the 27th international conference extended abstracts on Human factors in computing systems. pp. 2763-2766ACM, Boston, MA, USA (2009).
30. Roto, V., Väänänen-Vainio-Mattila, K., Law, E., Vermeeren, A.: User Experience Evaluation Methods in Product Development (UXEM'09). Human-Computer Interaction – INTERACT 2009. pp. 981-982 (2009).
31. Väänänen-Vainio-Mattila, K., Roto, V., Hassenzahl, M.: Now let's do it in practice: user experience evaluation methods in product development. CHI '08 extended abstracts on Human factors in computing systems. pp. 3961-3964ACM, Florence, Italy (2008).
32. Law, E., Roto, V., Hassenzahl, M., Vermeeren, A., Kort, J.: Understanding, scoping and defining user experience: a survey approach. Proceedings of the 27th international conference on Human factors in computing systems. pp. 719-728ACM Press, Boston, MA, USA (2009).
33. Roto, V., Vaananen-Vainio-Mattila, Law, E., Vermeeren, A.: User Experience Evaluation Methods in Product Development (UXEM'09). Lecture Notes In Computer Science, Vol. 5727. pp. 981-982ACM Press (2009).
34. Law, E., Roto, V., Vermeeren, A.P., Kort, J., Hassenzahl, M.: Towards a shared definition of user experience. CHI '08 extended abstracts on Human factors in computing systems. pp. 2395-2398ACM, Florence, Italy (2008).
35. Hassenzahl, M.: The thing and I: understanding the relationship between user and product. Funology: from usability to enjoyment. pp. 31-42Kluwer Academic Publishers (2004).
36. Cockton, G.: Designing worth---connecting preferred means to desired ends. interactions. 15, 54-57 (2008).
37. Cockton, G.: Value-centred HCI. Proceedings of the third Nordic conference on Human-computer interaction. pp. 149-160ACM, Tampere, Finland (2004).
38. Cockton, G.: From quality in use to value in the world. CHI '04 extended abstracts on Human factors in computing systems. pp. 1287-1290ACM, Vienna, Austria (2004).
39. International Organization for Standardization: ISO 9241-210:2010 - Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems, http://www.iso.org/iso/catalogue_detail.htm?csnumber=52075.
40. Hassenzahl, M.: Aesthetics in interactive products: correlates and consequences of beauty. In: Schifferstein, H. and Hekkert, P. (eds.) Product Experience. pp. 287-302Elsevier, New York NY (2008).
41. Kelly, G.: The psychology of personal constructs. Norton, New York (1955).
42. Fishbein, M., Ajzen, I.: Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Addison-Wesley, MA (1975).
43. Sluis-Thiescheffer, W., Bekker, T., Eggen, B.: Comparing early design methods for children. Proceedings of the 6th international conference on Interaction design and children. pp. 17-24ACM, Aalborg, Denmark (2007).
44. Druin, A.: Cooperative inquiry: developing new technologies for children with children. Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit. pp. 592-599ACM, Pittsburgh, Pennsylvania, United States (1999).

45. Sluis-Theischeffer, W., Bekker, T., Eggen, B.: Adding User Creativity to the UX toolbox: Exploring the use of Creative UX methods. Presented at the 13th CHI Netherlands, The Netherlands (2009).
46. Barendregt, W.: Evaluating fun and usability in computer games with children. Technische Universiteit Eindhoven (2006).
47. Ericsson, K.A., Simon, H.A.: Protocol Analysis: Verbal reports as data. The MIT Press, Cambridge, MA, US (1993).
48. Als, B.S., Jensen, J.J., Skov, M.B.: Comparison of think-aloud and constructive interaction in usability testing with children. Proceeding of the 2005 conference on Interaction design and children - IDC '05. pp. 9-16, Boulder, Colorado (2005).
49. Höysniemi, J., Hämäläinen, P., Turkki, L.: Using peer tutoring in evaluating the usability of a physically interactive computer game with children. Interacting with Computers. 15, 203-225 (2003).
50. Kesteren, I.E.,H.V., Bekker, M.M., Vermeeren, A.P.O.S., Lloyd, P.A.: Assessing usability evaluation methods on their effectiveness to elicit verbal comments from children subjects. Proceedings of the 2003 conference on Interaction design and children. pp. 41-49ACM, Preston, England (2003).
51. Read, J.C., Markopoulos, P.: LIFELONG INTERACTIONS. Understanding children's interactions. interactions. 15, 26 (2008).
52. Read, J.C.: Validating the Fun Toolkit: an instrument for measuring children’s opinions of technology. Cognition Technology and Work. 10, 119-128 (2008).
53. ISO 9241-210:2010 - Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems, http://www.iso.org/iso/catalogue_detail.htm?csnumber=52075.
54. Vanden Abeele, V., Zaman, B., Vanden Abeele, M.: The Unlikeability of a Cuddly Toy Interface: An Experimental Study of Preschoolers' Likeability and Usability of a 3D Game Played with a Cuddly Toy Versus a Keyboard. Presented at the 2n International Conference on Fun and Games, Eindhoven, The Netherlands (2008).
55. Lemish, D.: Preschoolers, media impact on developmental needs of. In: Arnett, J.J. (ed.) Encyclopaedia of Children, Adolescents, and the Media. pp. 407-410Sage, Thousand Oaks (2007).
56. Rice, F.: Human development : a life-span approach. Prentice Hall, Upper Saddle River N.J. (1998).
57. Bukatko, D.: Child development : a thematic approach. Houghton Mifflin, Boston (2001).
58. Irwin, L.G.: Interviewing Young Children: Explicating Our Practices and Dilemmas. Qualitative Health Research. 15, 821-831 (2005).
59. Bretherton, I., Beeghly, M.: Talking about internal states: The acquisition of an explicit theory of mind. Developmental Psychology. 18, 906-921 (1982).
60. Dunn, J., Bretherton, I., Munn, P.: Conversations About Feeling States Between Mothers and Their Young Children. Developmental Psychology. 23, 132-139 (1987).
61. Hickling, A., Wellman, H.: The Emergence of Children's Causal Explanations and Theories: Evidence from Everyday Conversation. Developmental Psychology. 37, 668-683 (2001).
62. Miller, P., Aloise, P.: Young children’s understanding of the psychological causes of behavior: a review. Child Development. 60, 257-285 (1989).
63. Berk, L.: Development through the lifespan, third edition. Allyn & Bacon, Boston (1999).

64. Bidell, T., Fisher, K.: Beyond the stage debate: Action, structure, and variability in Piagetian theory and research. In: Sternberg, R. and Berg, C. (eds.) Intellectual Development. pp. 100-140Cambridge University Press, Cambridge, England (1992).
65. Johnson, M.P., Wilson, A., Blumberg, B., Kline, C., Bobick, A.: Sympathetic interfaces: using a plush toy to direct synthetic characters. Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit. pp. 152-158ACM, Pittsburgh, Pennsylvania, United States (1999).
66. Paiva, A., Andersson, G., Höök, K., Mourão, D., Costa, M., Martinho, C.: SenToy in FantasyA: Designing an Affective Sympathetic Interface to a Computer Game. Personal Ubiquitous Comput. 6, 378-389 (2002).
67. Vanden Abeele, V., Zaman, B., Vanden Abeele, M.: The Unlikeability of a Cuddly Toy Interface: An Experimental Study of Preschoolers' Likeability and Usability of a 3D Game Played with a Cuddly Toy Versus a Keyboard. Fun and Games. pp. 118-131 (2008).
68. Zaman, B., Vanden Abeele, V.: How to Measure the Likeability of Tangible Interaction with Preschoolers. Proceedings of CHI.NL. pp. pp. 57-59Infotec Nederland BV Woerden, Eindhoven, The Netherlands (2007).
69. Andersson, G., Höök, K., Mourão, D., Paiva, A., Costa, M.: Using a Wizard of Oz study to inform the design of SenToy. Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques. pp. 349-355ACM, London, England (2002).
70. Höysniemi, J., Hämäläinen, P., Turkki, L.: Wizard of Oz prototyping of computer vision based action games for children. Proceedings of the 2004 conference on Interaction design and children: building a community. pp. 27-34ACM, Maryland (2004).
71. Zaman, B.: Introducing a Pairwise Comparison Scale for UX Evaluations with Preschoolers. Human-Computer Interaction – INTERACT 2009. pp. 634-637 (2009).
72. LadderUX, on line and freely available software tool for UX laddering, www.ladderux.org.
73. Pieters, R., Baumgartner, H., Allen, D.: A means-end chain approach to consumer goal structures. International Journal of Research in Marketing. 12, 227-244 (1995).
74. Reynolds, T., Gutman, J.: Laddering theory, method, analysis, and interpretation. In: Reynolds, T. and Olsen, J. (eds.) Understanding consumer decision making : the means-end approach to marketing and advertising strategy. pp. 25-52Lawrence Erlbaum, Mahwah N.J. ;;London (2001).
75. Vanden Abeele, V., Zaman, B.: Laddering the user experience! User Experience Methods, Interact 2009, Uppsala, Sweden. , Uppsala, Sweden (2009).
76. Grunert, K.G., Beckmann, S.C., Sorensen, E.: Means-end chains and laddering: an inventory of problems and an agenda for research. Understanding Consumer Decision-Making: The Means-end Approach to Marketing and Advertising Strategy, pp. pp.63-90.Lawrence Earlbaum, Mahwah, NJ (1996).
77. Reynolds, T.J., Olson, J.C.: Understanding Consumer Decision Making. (2001).
78. Zaman, B.: Introducing contextual laddering to evaluate the likeability of games with children. Cognition, Technology & Work. 10, 107-117 (2007).

Figures

Tables

Please note that this is the final draft post-refereeing print. For the publisher's version and for citation purposes, please use: Vanden Abeele, V., Zaman, B., De Grooff, D. (2012). User eXperience Laddering with Preschoolers: Unveiling Attributes and Benefits of Cuddly Toy Interfaces. Personal and Ubiquitous Computing, 16 (4), 451-465.