Laddering with Young Children in User eXperience Evaluations: Theoretical Groundings and a Practical Case

Bieke Zaman
CUO/IBBT
Katholieke Universiteit Leuven,
Parkstraat 45, bus 3605,
3000 Leuven, Belgium
+32 16 32 36 51
Bieke.Zaman@soc.kuleuven.be

Vero Vanden Abeele
Vesaliusstraat 13
3000 Leuven, Belgium
+32 16 30 10 30
Vero.Vanden.Abeele@groept.be

ABSTRACT
In this paper, we investigate the usefulness and feasibility of Laddering with young children in User eXperience evaluations. We start with a revision of theoretical literature and guidelines. Developmental literature suggests that children aged two to seven years old have the cognitive capabilities to perform as Laddering interviewees. Next, we put these findings to the test via a practical case. The results of our case study demonstrate that only the older children, aged five years and older, were able to construct meaningful ladders. As for the type of ladders created, our results are in line with literature; children are inclined to attribute external reasons to product preferences rather than internal reasons, and consequently create ladders of attributes and consequences, not reaching for values.

Categories and Subject Descriptors
H.5.3 [User Interfaces]: Evaluation/methodology

General Terms
Measurement, Design

Keywords
Laddering, young children, UX, User eXperience, evaluation methods

INTRODUCTION
User studies in which children try out new technologies can be very interesting and informing. Yet, in order to suggest design improvements regarding usability or User eXperience, we should be careful with drawing conclusions from children’s enthusiastic play behaviours and reactions. Researchers should always critically question whether children’s enthusiasm is really caused by positive User eXperiences and not for instance by the fact that in general people are more likely to report positive emotions to products than negative [33].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

It may also be that the enthusiasm can be attributed to the specific research situation [28]. Children in general enjoy the situation in which adults observe them while playing, receiving undivided attention and sometimes even rewards [25]. Without proper evaluation methods or research set-ups, researchers do not exactly know what the enthusiasm accounts for. This fuzzy accountability has its implications for the design process, e.g. it limits the detection of design attributes that cause problems and should be improved, or it hampers the correct estimation of successful adoption of the product.

In the early design phases, researchers rely on creative methods to inform the design and get insight into children’s likes and dislikes [34]. Examples of these creative methods or techniques are mood boards, cooperative inquiry [15], drawings, sticky note critiquing, Mission from Mars [14], or low-tech prototyping. In the middle and late design phases, typically some test materials are available, be it sketches, models, scenarios, mock-ups or prototypes [22,34]. Therefore user-based evaluative methods become prominent, such as the Problem Identification Pictures Cards (PIPC) method [2], (spontaneous) thinking-aloud [16], constructive interaction [1], peer-tutoring, post-task interviews [25], Self Assessment Manikin (SAM) [6] and the Fun Toolkit –Smileyometer, Funometer, Again-Again meter- [28]. Whereas in the beginning of a design process, the main research aims deal 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 key attributes to improve towards an optimal User eXperience [22].

In this paper, we will add Laddering, a specific interview and data analysis technique, as complementary to these evaluative methods. We argue that Laddering has much potential. First, it allows for revealing more in-depth information than pictorial representational scales such as SAM [6], PIPC [2] or the Smileyometer [28]. Further, Laddering not only reveals preferences or lists positive and
negative issues. It goes further and digs into the reasons why a product was liked or disliked, pointing out to concrete design issues that should be solved or improved. Last but not least, the Laddering method also compensates for the shortcomings of think-aloud or constructive interaction protocols that are likely to cause a cognitive burden. Laddering similarly requires experiencing a product and commenting on that experience [5], but in a gated and incremental procedure that aligns well with the cognitive limitations of young children.

We will begin this paper by explaining the Laddering method, entailing both a qualitative interview method and a quantitative data analysis method. Next, we will describe young children’s relevant cognitive and behavioural capabilities in relation to the Laddering interview context and argue why Laddering is especially suited as an evaluation method for young children. In the second part of the paper, we will put the theory into practice and demonstrate the Laddering interview method via a practical case. Finally, we will check and discuss the results of a Laddering case study against the literature findings and explore ideas for further work.

I. THEORETICAL GROUNDINGS

The Laddering method

Laddering originated in consumer research and relies heavily on Means-End Theory as proposed by Gutman [19]. 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 have certain desired consequences as beneficial towards their individual values. The common generic means-end chain, therefore, consists of attributes (A), consequences (C) and values (V).

Attributes \(\rightarrow\) Consequences \(\rightarrow\) Values

Although firstly conceptualized in the sixties, the term Laddering and the wider acceptance of the technique only broke through in the eighties. In 2004, Cockton advocated the deployment of customer psychology and the associated means-end chains in HCI for studying User eXperience [10]. In general, Laddering is most commonly known as a specific in-depth interviewing technique used in qualitative research [18]. However, one may not forget that full Laddering entails both a qualitative interview process and a specific procedure for quantitative data analysis.

Laddering as interview technique

Laddering starts by eliciting distinctions in products or product classes [29]. The techniques that are used for this distinction elicitation can be chosen by the interviewer. One relevant technique consists in asking which product is preferred and why. This technique is referred to as the preference-consumption differences technique or the ranking technique [29,35]. In explaining these product (class) preferences, the interviewee is first prompted to identify those salient attributes (A) that explain the preference or distinguish the preferred product from the alternatives. Then, the interviewer will try to reveal the explanatory consequences (C) and values (V) for the product preference and thus create a means-end chain or an A-C-V sequence or ladder.

Imagine for instance that a child expresses his preference for game x as opposed to game y (preference), because the first is played with a steering wheel as opposed to the latter which is controlled by a joystick (salient attributes that explain the preference). After this initial attribute elicitation phase, the interviewer will try to reveal the child’s product attribute related ladder by asking questions like “Why is this [attribute] important to you?” or “What does it mean to you?”. In other words, the child is asked to motivate his attribute selection by explaining the related anticipated and favoured consequences. In our example, the interviewer may ask “why do you find it important that a game is played with a steering wheel?” or “What does it mean to you?”.

The interviewer, however, does not rest at the consequences level and keeps prompting “Why?”. 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. This way, theoretically, the reasons (consequences) why certain attributes are important will be revealed, followed by an expression of how these consequences serve personal values. To illustrate, the interviewer may reveal that playing with the challenging steering wheel (Attribute) results in gaining a significant amount of points in the game (Consequence) which is important for the child who likes to be a winner and perform best (Value).

Laddering as data treatment

The Laddering data analysis process typically involves the following two phases, a qualitative and a quantitative phase. In the qualitative phase, the Laddering interviews are transcribed and meaningful core elements (i.e. attributes, consequences, values) are coded. Once that the core elements are defined and labelled, the individual ladders can be decomposed based on these codes. Then, in the quantitative phase, the data from all individual ladders are summarized and the relations between all elements quantified in terms of direct and indirect links. In order to focus on the most relevant data and reveal dominant patterns of thinking and perceptions at the aggregate level, the researcher has to decide upon a cut-off level of direct or indirect links. Links between elements that fall beneath cut-off levels are ignored because they are too marginally related; those that reach that predefined level are kept.
Finally, an overview of the dominant links between attributes, consequences and values—the means-end chains at the aggregate level—are typically visualized in a map, called a Hierarchical Value Map (HVM) (see figure 3). The HVM obtained through the Laddering procedure offers valuable information. It can serve as a basis for: (1) comparing different scenarios, prototypes, designs, brands or products; (2) understanding the impact of design features on user appreciation; and (3) consequently result in informed design adjustments.

**Young children and Laddering**

Laddering has gained substantial interest when researching adults and product preferences [31]. However, before any theoretical assumptions can be made about the feasibility and relevance of Laddering for young participants, we need to comprehend the typical abilities and skills of this young age group. Many theorists have categorized children according to certain developmental stages. Children belonging to one related age group can be characterized in terms of relatively similar physical growth, motor development, cognitive or intellectual development, social or emotional development and language. Jean Piaget labeled preschoolers’ behavioural and cognitive characteristics as preoperational, encompassing children from two to seven years old. In this paper, we will deal with this preoperational age group of preschoolers, who cannot yet read or write and did not yet start elementary school.

One of the most prominent capabilities that children from the preoperational stage acquire is language [24]. Preschoolers’ language has reached a sufficient level for the adequate use of words to describe actions or represent objects [8,32]. The same goes for their pragmatics, or “children’s practical ability to use language to communicate with others in a variety of social contexts” [32:168], and preschoolers’ ability to deal with the complexities of open-ended questions [23]. 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. This implies that they do not need the object or the experience to be recent and apparent to be able to discuss these [24]. Nevertheless, young children’s language and pragmatics capabilities only prove to be efficient in low-demanding situations in which they can see the interviewer’s reactions or rely on typical conversational aids as gestures, verbal prompts or cues and objects to talk about [3]. 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 [32]. In the case of Laddering, we should therefore ensure that the technologies evaluated are meaningful objects used in a familiar context-of-use.

In order to grasp children’s perceptions of their world, it may be that traditional questioning techniques fall short with child respondents. Often, the researcher should even engage in children’s activities. That is what Corsaro [12:28] calls the reactive strategy in which adults are acting in a *least adult* way in their encounters with children. Similarly to the way children behave amongst peers, adults should approach children in a reactive way, e.g. in initiating contact, or by actively entering their natural environment and engaging in their activities. Taking Corsaro’s advice into account and projecting it onto the Laddering situation, the interviewer can for instance build rapport with the child by participating in his/her activities first and wait to ask questions about the relevant play activities until he/she was approached by the child him- or herself.

Apart from their language and pragmatics, young children’s awareness of mental states and their understanding of psychological causes are necessary 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 [7] as well as their ability to think of events and behaviours in terms of causal relations [21]. Miller & Aloise [26] 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 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* [26]. Miller & Aloise [26:259] 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 [3,8]. Often forgotten, Miller & Aloise [26] also stressed that most causal expressions of young children are even reflected upon things that they intend to do next instead of things that they have just experienced.

If children are asked to reflect upon their User eXperiences and the corresponding object preferences, then two types of reasons can be given. On the one hand, children may explain their preference in terms of *external causes*, related to the characteristics of the object entity (e.g. “because it has many challenging obstacles.”). On the other hand, *internal reasons* may be given to explain their choice, related to the characteristics of the person who made the choice (e.g. “I like overcoming challenging situations”) [26]. Because of their cognitive limitations, young children are inclined to attribute external reasons over internal reasons to their preferences. First, young children have a tendency to rely on stable and consistent causes to explain things in life, hereby *emphasizing external causes* as explanations of their preference. Secondly, many children from the preoperational phase do not yet have sufficient capabilities to reason about abstract concepts which may be necessary when dealing with internal reasons.
The research set-up of a Laddering interview similarly promotes the revelation of external causes before addressing internal causes. A Laddering interview explicitly deals with product differences; one cannot start a discussion without referring to the product attributes in question [26]. Further, in a situation in which the interviewer often specifies the why-probing question into a question like for instance “Why did you prefer that game?”, the attention is first drawn towards the (external) characteristics of the product. It is only when probing further into the reasons for product preferences that the researcher tries to reveal internal reasons. Nevertheless, in this process of going from concrete product attributes (external) to more abstract (internal) values, we will run into children’s cognitive limitations. When probing into internal reasons, young children will tend to evaluate the self’s characteristics (attributes, abilities and attitudes) in a concrete, observable and –often exaggerated- positive way (“I am happy when I play that game”, “I am the best in playing that game”) [3]. In other words, although mental state terms are used appropriately by young children, these references do not reflect a true stable internal characteristic that can be generalized to many situations, as with older children or adults [26]. Thus, young children are inclined to attribute external reasons to their preferences, or else describe their internal preferences in such a context-dependent way that it cannot always be generalized to other situations; neither does it reflect the child’s general needs or values.

To conclude, children from the preoperational stage already show cognitive and behavioural capabilities that are needed to be adequate respondents in a Laddering interview.

On a final note, although children share many similar capabilities, there is more variation in the preoperational stage than may be apparent from Piaget’s stage theory [4]. In other words, although we refer to the typical achieved intellectual and language developmental capabilities of the preoperational children, we acknowledge that individual and cultural aspects account for significant in-group differences.

UX evaluations and Laddering

Basing ourselves on the International Standard Organisation’s definition of User eXperience (UX) [22] and taking into account the insights from other academic (UX) experts, we adhere to the following operational definition of UX as “a person’s perceptions and responses that result from the use or anticipated use of a product, system or service” [20]. According to this definition, User eXperience refers to an individually constructed, evaluative (overall or attribute related) judgment of and response to the perceived actual and recent experience of interacting with a product or system in a particular context-of-use.

In order to understand this definition better, we find Hassenzahl’s [20] explanation of how judgmental responses on product use (e.g. in terms of preferences) are built, especially useful. More particularly, Hassenzahl [20] brought two key explanatory factors into prominence: the impact of a) perceived product attributes and b) personal standards or expectations. He thereby explained that people are likely to judge products against other relevant products by considering the specific product features that make a difference. Hence, in order to explain preferences, people are likely to relate to experiences with other relevant products, systems or services and consider how these other relevant systems or services (means) can meet their personal standards and expectations (ends).

Laddering responses as attribute related explanations of overall preference judgements

From a methodological point of view, the Laddering interview technique addresses the two key factors that are decisive for preference judgments. The basic premise is that several options are to be compared and judged on importance using personal standards or expectations. More particularly, Laddering relies on comparisons concerning perceived and meaningful differences between products [29]. In the case of Laddering with young children, the interviewer asks the child to decide upon the most important or preferred alternative. Secondly, Laddering goes further than the overall judgment alone. The preferred option serves as the basis for a probing session during which the researcher asks about the reasons why that product (attribute) was preferred over the other, hereby especially probing into the judgments of specific product attributes according to personal standards or expectations.

Attribute-eliciting versus value-eliciting approaches

In general, we can distinguish two approaches towards means-end analysis, which has pertinent methodological implications. Some researchers such as [11] focus on the measurement of personal standards and expectations (ends or values) first. In contrast, others such as [20] prefer to focus on product attributes (means) first. As for both approaches, the analysis of the results remains the same and is aimed at revealing attribute-consequences-value chains, also known as means-end chains.

We recommend attribute-elicitation measurement techniques, rather than value-elicitation techniques for UX evaluations with young children, for the following two reasons. Firstly, as aforementioned, children’s competency to talk about abstract values is rather limited. As a consequence, direct instruments to access values such as closed-value surveys do not seem feasible. Even adults are likely to have difficulties verbalizing what is really important to them when thinking in terms of values [11]. Laddering compensates for young children’s cognitive limitations. Young children indeed prefer and are more able to rely upon external causes in favour of internal causes or personal values in their explanations of events. Secondly, we argue that in a HCI or child-centered design context, the aim of many evaluation projects is to reveal those product features that can be improved in order to design for optimal User eXperiences.
In other words, the design’s locus of attention is primarily based at product attributes. A designer needs practical recommendations to improve the product’s attributes. The motivations for these improvements, then refer to the corresponding consequences or values that can be enhanced.

Methodological implications for UX Laddering with young children

Laddering in context

Although we define UX as individually constructed perceptions and responses, we guard ourselves from neglecting contextual factors. Within the child-computer interaction community, it has been strongly promoted to invite children to a test setting that they are familiar with to make them feel at ease [25]. Moreover, as we discussed earlier, children’s cognitive capabilities perform best in situations that are meaningful and relevant to them. Consequently, research at school or at home is preferred over laboratory settings. Nevertheless, researchers should not only prefer a test location that answers the condition of familiarity, but also the condition of ecological validity. That is why researchers should ideally look for a test location that matches the context where children are likely to encounter the technology that is being evaluated [25].

Reynolds and Gutman bring this issue back to Laddering: “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 behavior provides a more meaningful context for Laddering to proceed. People do not use or consume products in general; they do so in particular contexts [29:16].”

In previous research, we found that ensuring recent experiences e.g. by an exploratory use of the object under test, facilitates the Laddering interview with children [41]. When the explorative play precedes the interview, contextual clues such as language and non-verbal signs can be used as a basis to formulate questions that are tailored to the child’s individual situation. Markopoulos et al. [25] came to the same conclusions when discussing the advantages of complementing observations of technology experiences with verbal reports. By obtaining opinions on UX in the context of object interaction, the verbal responses can be related to the specific aspects of this concrete previous experience [25]. Young children need to rely on typical conversational aids such as gestures, verbal prompts and objects to talk about to maximize their abilities to use language in communications [3] and deal with open-ended questions [23]. These conditions can be met if the object in question stays in front of the child during the Laddering interview. The object makes the interview more concrete and gives the child the possibility to actively show what he/she wants to explain. Showing things might be easier for children than explaining and thus ‘translating’ into words.

As for the show-me strategy [13] or what we call action rather than reflection-strategy, we take Corsaro’s [12] advice into account and promote the engagement in activities with children to complement the data obtained through observations and verbal reports. If it is hard to receive any information from the child, the researcher should ask the child to show a certain attribute or to mimic a specific play. Darbyshire et al. [13] reported on the show-me strategy’s positive effects such as the stimulation of children’s thinking and discussion. Action and play enhance children’s motivation and memory; and they allow the researcher to ground his/her questions in the here and now. In other words, in those situations when children have difficulties to reflect upon their technology experiences, one should motivate them towards action (‘show, tell & play together’) since action triggers reflection. Moreover, this interactivity is likely to enhance the child’s motivation, breaking through the rather conventional sit-down interview situation [17]. In order to meet these conditions further, a great responsibility falls on the adult interviewer who has to create a child-appropriate interview setting. In this process, the adult’s role slightly changes from an interviewer to a participant observer.

Child-friendly interview setting

In order to ensure a child-friendly interview setting we should think of ways in which the interviewer can take a least adult position [12]. In this context, Yarrow [40] suggests to assign the child an expert role and Darbyshire et al [13] propose the interested idiot research strategy. As for interested idiot strategy, the adult explicitly leaves the position of the all-knowing adult for an ignorant person, transforming the child into the expert in question. A concise introduction to the interview is needed to motivate and prepare the child for the role of the interviewer as a totally ignorant adult. In a Laddering context, the researcher should explain the test child that he/she will be asking the same -obviously and possibly even stupid- why-question over and over again [30]. During the interview, the child can be encouraged by labeled praise (e.g. “You are doing a good job by telling me what you think”) [39]. Further, one should try to ask mainly clear and open-ended questions since they elicit more information per question [39]. When probing for a deeper understanding of the User eXperience, open-ended questions are preferred since they are likely to result in more objective answers than closed-ended questions, especially when there is a strong possibility of socially desirable answers [36].

Apart from the specific interview or probing strategies mentioned above, we should not forget to mention general tips for interviewing children. E.g. the spoken language should be adapted to the child. This can be realized through an appropriate vocabulary, sentence structure, the use of silence or points of interest [39]. In order to avoid children mentioning information to please the adult, one should explicitly allow “don’t know”-answers [9].
The interviewer should also stress that every answer is a correct answer and that there are no wrong answers [30]. Furthermore, Laddering interviews with children should remain short. Contrary to Laddering interviews with adults which can take over an hour, child-Laddering situations should demand less time and less concentration. Especially when an exploration phase precedes the Laddering interview, during which the child is intensively learning to use the product. It is not only a matter of concentration, but also a matter of willingness to participate. When the interview lasts too long, children become bored. The interview situation is indeed rather tedious compared to the previous explorative play experience. Young children are not likely to have had any experience with the typical sit-down research interviews [17:112] and consequently, might not feel very comfortable. Nevertheless, young children can be very keen to participate in an interview. They are especially pleased to be treated as someone who has something useful to contribute [9] as long as the conditions of the interview are adapted to their interests and capabilities.

As a final guideline, there are four strategies that help to overcome difficult Laddering scenarios. If the child cannot explain what (s)he actually liked about the preferred technology, then the researcher can first try to rephrase the question into a negative sentence by inquiring into the reasons why the child did not like the other technology compared to the preferred one. Secondly, the researcher might also ask what would happen if a preferred attribute is taken away or a consequence not delivered [30]. Another solution that can be used in a situation in which the child cannot explain his/her choice, is to rephrase the question by making it more concrete (e.g. giving a situational example or referring to contextual clues) [30]. Last but not least, in a situation where the interviewer him/herself encounters difficulties to think of appropriate questions, Walker et al. [39] propose to paraphrase what the child has said. This helps to organize the child’s thoughts and feelings and results in increased verbal interchange [39].

II. CASE STUDY

In this second part of the paper, we will test and discuss the feasibility of Laddering with young children via a practical case. TOEWIE 2.0 is an on-going research project aiming at designing tangible cuddly toy controllers for 3D games for preschoolers [38]. A discussion of the entire project is beyond the scope of this paper. In this paper, we will particularly focus on the project’s design phase where evaluation was needed to benchmark three different prototypes with respect to their User eXperience: a bird game, a kangaroo game and a penguin game (see figure 1). The aim of the evaluation was to decide upon which prototype was to be developed further and what attributes could be improved.

![Figure 1. An overview of the three different cuddly interfaces and the game worlds.](image)

**Research aims**

With regards to Laddering with children, we aimed at answering the following research questions. First, we wondered whether our preschoolers would be able to express a relatively consistent preference for one of the three prototypes (RQ1). Secondly, we were interested in whether our young interviewees would mention attributes, consequences and or values -and thus constitute full ladders- as explorative for their game/interface preference (RQ2). Our last research question dealt with the number of ladders and accompanying elements children created in general, and what the impact of age was for the numbers of ladders and or elements in particular (RQ3).

**Method**

**Participants & Location**

All interviews took place at an elementary school, where during holidays, day camps were organized for kids of the community. The testing happened during these summer day camps, on three consecutive days. 46 preschoolers that were visiting the day camps participated in the Laddering study (M=67.54 months, SD = 13.54 months). The youngest participant was 33 months old; the oldest was 86 months 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 steered the interaction, both remained in the background and did not intervene in the interview.

**Procedure**

As was suggested by literature, we foresaw time for an explorative game play preceding the Laddering interview. As the child entered the room, he or she was invited to start playing with the three cuddly toy interfaces. The order of play with the prototypes was counterbalanced to avoid order effects. After playing with the three different prototypes, the child was invited to sit in front of the three different prototypes (see figure 2). Next, the preferred choice was first elicited via the This-or-That method [42] that relies upon five questions to infer which prototype was preferred most by the preschooler. Four out of five questions were positively formulated (“show me which game... was most fun”, “...would you like to receive as gift”, “...would you like to play again”, and “...would you like to take home?”). One question was negatively formulated (“Show me which game you found a little bit stupid?”).
After the This-or-That questions, the child was prompted once more to identify his/her preferred choice upon the sentence: “I’m sorry, I have forgotten, can you show me once more which game you found most fun?”. The answer to this last question -from now on referred to as ‘Laddering preference’- was the basis for the Laddering interview during which the facilitator probed for the reasons why that identified prototype was preferred.

In order to investigate children’s preference consistency, we first analyzed the relation between the Laddering preference on the one hand and the expressed preference via the This-or-That questions on the other hand. We found that in general children were consistent in their preferences. If we focus on the cases in which the kangaroo game was preferred over the penguin or bird game, we found that there were more consistencies than we would expect by accident ($L^2(2)=(8.5)$, $p<0.05$). The same holds true when we focus at the cases in which the penguin game was preferred over the bird game or the kangaroo game ($L^2(2)=(42.9)$, $p<0.01$) and the cases in which the bird game was preferred over the penguin or kangaroo game ($L^2(2)=(25.2)$, $p<0.01$). Secondly, we compared the preference uttered via the This-or-That questions and the free play option, and again we found consistent patterns of relations that correspond more than one would expect by accident (when focusing on penguin selection versus the other games: $L^2(4)=(30.5)$, $p<0.01$; when focusing on the kangaroo preference versus preference for one of the other games: $L^2(4)=(25.7)$, $p<0.01$ or when focusing at the preference for the bird game versus one of the other two games $L^2(4)=(25.4)$, $p<0.01$). As for our first research question, we can thus conclude that in general, children expressed a relatively consistent preference for one of the three games, both verbally and behaviourally. As a consequence, we may rely upon children’s answers on the question “which game did you find most fun” as the start of a Laddering interview.

Results
RQ1. Consistency of preference
Our first research question was aimed at finding out whether our preschoolers would be able to express a relatively consistent preference. To do this, we had to calculate Likelihood Ratio values, which are an alternative to Pearson’s Chi Squares, to be used when small data sets are involved and when expected values reach frequencies below five. Moreover, instead of working with 3 x 3 contingency tables in which the three prototypes are compared simultaneously, we created dummy tables by focusing at the preference of one game versus the preference of ‘one of the two other games’. This way, we could increase the expected cell frequencies.

Figure 2. An overview of the interview setting.

If a child remained silent, repeated him/herself, or indicated in any other way that it was impossible to answer, we finished the interview. At the end of the interview, children were rewarded for their good participation by receiving the option to play one of the three games again. This free play choice was assumed to be correlated with the most preferred game revealed via the This-or-That questions on the one hand and the expressed ‘Laddering preference’ on the other hand.

Qualitative data analysis
All interviews were recorded and transcribed ad verbatim. Next, the transcriptions of the interviews were analyzed by both authors independently, to establish the core elements (i.e. the dominant attributes, consequences and values) from the interviews. These two lists of elements were compared to each other, and refined into one set of elements that both researchers agreed upon. 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 0.77 ($p <0.001$), an acceptable level of agreement between coders.

Results
RQ2. Complete ladders
The second research question focused on whether our young interviewees would express complete ladders. Our case study results show that children spontaneously mentioned attributes and consequences (see table 1).

<table>
<thead>
<tr>
<th>Concrete Attributes</th>
<th>Functional consequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>phys. flapping</td>
<td>exper. extraordinary things</td>
</tr>
<tr>
<td>phys. wiggling</td>
<td>being easier/faster</td>
</tr>
<tr>
<td>phys. hopping</td>
<td>novel/different</td>
</tr>
<tr>
<td>flying</td>
<td>similar/recognition</td>
</tr>
<tr>
<td>walking</td>
<td>beautiful/cool/nice to watch</td>
</tr>
<tr>
<td>jumping</td>
<td></td>
</tr>
<tr>
<td>aest. tumble</td>
<td>Psycho-social consequences</td>
</tr>
<tr>
<td>aest. topple</td>
<td>immersion/fantasy</td>
</tr>
<tr>
<td>aest split</td>
<td>funny/humor/laughing</td>
</tr>
<tr>
<td>apples</td>
<td>social interaction</td>
</tr>
<tr>
<td>fish</td>
<td>mastery</td>
</tr>
<tr>
<td>graphical details</td>
<td></td>
</tr>
<tr>
<td>Abstract attributes</td>
<td>Terminal Value (Conceptual)</td>
</tr>
<tr>
<td>soft/cute/cuddly</td>
<td>Fun</td>
</tr>
<tr>
<td>healthy/tasty</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. List of attributes and consequences
Although we could easily distinguish 14 attributes (some referring to actual movements the children had to make, some to the cuddly toy itself, some referring to the in-game animations and objects), we could only distinguish eight different consequences and none of the children mentioned values to explain their product preference.

When investigating the interviews, children could often point at attributes but failed to climb up the ladder and did not address any consequences. Of the 46 children that participated, nine children could not point out any specific attributes; they only expressed a preference for a specific prototype. 13 children mentioned attributes but failed to list consequences (see example below).

**Researcher:** I forgot again, which one did you like most?

(child points at the penguin)

**Oh yes, why did you like that one the most?**

"because it has fishies" [Attribute]

**Why do like that?**

"because they are in the water" [A]

Ah I see, **why do you like that fishes are in the water?**

(... silence)

**What else did you like about the penguin**

That I could lift his wings [A]

**Oh yes, that's right, why did you like that?**

(... silence)

Yet, 24 children effectively did mention both attributes and consequences as illustrated below. In the example, it is remarkable that when the interviewer kept prompting why, the child started pointing out attributes again instead of climbing up the ladder towards more consequences or even values. This climbing down the ladder was something that happened in a number of interviews, as children seem to interpret the repeated why-questions of the interviewer as indicative of not having understood the attributes. This research finding is in line with the emphasis on external causes as reasons for preference [26]

**Researcher:** I forgot again, which one did you like most?

Child stands up and points at the bird

**Oh yes, the bird, can you tell me why?**

"Because I like it"

**Ah, because you like it, and can you tell me why?**

"The tumbling it makes" [A]

**Ah, the tumbling it makes, and why is that nice?**

"Because that is so funny" [C]

**Ah, and why do you like it when it is funny?**

"Because it is really cool that it can make a looping" [A]

**Ah, and why do you like that it makes a looping?**

"Because sometimes they fall on their bum and they do funny"[A-C]

**Ah, and why do you like that it is funny?**

"That they fall on their bum" [A]

Although none of the ladders attained the value level, we argue that due to the nature of the question: “Which did you find most fun?”, conceptually every ladder links from the prototype to the terminal value of ‘fun’. Therefore, even those children that only pointed to the prototype generated a ladder as well.

**RQ3. Numbers of ladders/elements and impact of age**

As for the last research question (RQ3), dealing with the number of ladders and elements and the impact of age, we found that in total 79 ladders and 223 elements were derived from the 46 preschoolers. On average, our participants generated 1,72 ladders and 2,81 elements. The goal of Laddering is not only to understand which prototype is preferred the most, but even more important, also to understand dominant means-end chains: which prototypes and related attributes lead to which consequences and values. This network of chains is typically graphed out in a Hierarchical Value Map (HVM). When generating enough ladders and elements, this decision can be done quantitatively by counting the number of direct and indirect links between elements. In this case, a cut-off level of three (to be roughly interpreted as whenever an element was linked three times to another element) resulted in the HVM below (see figure 3). The detailed discussion of the specific means-end chains is beyond the scope of this paper but can be found in [37]. Our results indicate that Laddering with children is possible and can yield insight in means-end chains that are valuable both for the evaluation of products and design suggestions for product improvements.

**Figure 3. The hierarchical value map.**

As yet mentioned, nine preschoolers 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), as compared to the mean age of the preschoolers producing ladders, M= 60.89 months (SD = 12.17). 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 positive correlation between age and number of elements was, r. 469 (p<0.01).
In sum, the results regarding our third research question show that it is possible to conduct Laddering interviews with preschoolers, but younger preschoolers are less likely to produce ladders and more likely to generate less elements within a ladder. If one aims for an interview that generates on average five elements per preschooler, linear regression analysis predicts an age of 59.57 months or almost five years ($r = 0.469$, $b_0=-1.314$, $b_1=0.106$, $p<0.01$).

**DISCUSSION**

Although developmental literature seems to suggest that children aged two to seven years old have the cognitive capabilities to perform as Laddering interviewees, the results of our case study suggest that only the older children, aged five years and older, were able to construct meaningful ladders. According to these figures, with a set of 45 preschoolers and 5 elements, the research study should in total provide 225 elements, which is close to the 223 elements we found in our case study, yielding an interesting Hierarchical Value Map. However, according to Reynolds and Olson [31], a traditional Laddering study should aim for approximately 500 elements. This number includes value elements as well. We do acknowledge that 223 elements resulted in the minimum cut-off level of three, (cut-off levels between three and five are recommended) in order to maintain approximately 66% of links at the aggregate level.

Further, we proposed Laddering as an individual technique to measure individual experiences. Future work may investigate whether inviting pairs of children simultaneously would benefit the Laddering interview situation and results. We might then question whether children would feel more comfortable and utter more valuable expressions, whether it would transform the data and whether it would be worth the effort of double recruiting and analysis [25,28]. Finally, future work may try out Laddering with children where the adult uses a hand puppet to administer the interview. Some researchers promote the use of a hand puppet in research with preschoolers [27] whilst others report on the artificiality of this kind of ‘mediated’ conversations [2].

**CONCLUSION**

In this paper, we began by theoretically questioning the sense and feasibility of performing Laddering interviews with preoperational children from two to seven years old in a UX evaluative context. Literature review revealed that children have relevant cognitive capabilities to function as adequate respondents in a Laddering interview as long as the interview is meaningful to them and occurs in a low-demanding, familiar setting. In our case study, we took the advices from literature into account, e.g. by having an explorative game play preceding the interview. Our results showed that Laddering is feasible, but only with the older preschoolers aged almost five years and older. We conclude that these older preschoolers are able to express ladders of concrete attributes and consequences, yielding insightful means-end chains.

**ACKNOWLEDGMENTS**

We like to thank the animators of the summer day camp Kattestroef and the community of Bertem, JITSO for granting us the permission to conduct this research on their play grounds. We like to thank master students Leonard Van Domselaer and Michiel Van Minnebruggen. We also like to thank he 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 finally Frederik Windey for the aid in the game development.

**REFERENCES**

34. Sluis-Theischemeier, W., Bekker, T., and Eggen, B. Adding User Creativity to the UX toolbox: Exploring the use of Creative UX methods. (2009), 131-140.
37. Vanden Abeele, V. and Zaman, B. Laddering as a User Experience Evaluation Method of Tangible Interactions with Preschoolers: a Case Study. Personal and Ubiquitous Computing, Special Issue on Tangibles for Children.