Which Robot Behavior Can Motivate Children to Tidy up Their Toys? Design and Evaluation of "Ranger"

Julia Fink
Séverin Lemaignan
Pierre Dillenbourg
Computer-Human Interaction in Learning and Instruction Laboratory (CHILI)
Ecole Polytechnique Fédérale de Lausanne (EPFL)
CH-1015 Lausanne, Switzerland
firstname.lastname@epfl.ch

Philippe Rétornaz
Florian Vaussard
Alain Berthoud
Francesco Mondada
Robotic Systems Laboratory (LSRO)
Ecole Polytechnique Fédérale de Lausanne (EPFL)
CH-1015 Lausanne, Switzerland
firstname.lastname@epfl.ch

Florian Wille
Karmen Franinović
Interaction Design (IAD)
Zürcher Hochschule der Künste (ZHdK)
CH-8005 Zürich, Switzerland
firstname.lastname@zhdk.ch

ABSTRACT

We present the design approach and evaluation of our prototype called "Ranger". Ranger is a robotic toy box that aims to motivate young children to tidy up their room. We evaluated Ranger in 14 families with 31 children (2-10 years) using the Wizard-of-Oz technique. This case study explores two different robot behaviors (proactive vs. reactive) and their impact on children's interaction with the robot and the tidying behavior. The analysis of the video recorded scenarios shows that the proactive robot tended to encourage more playful and explorative behavior in children, whereas the reactive robot triggered more tidying behavior. Our findings hold implications for the design of interactive robots for children, and may also serve as an example of evaluating an early version of a prototype in a real-world setting.

Figure 1: The robotic box "Ranger" aims to motivate children to tidy up their toys on the floor

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; I.2.9 [**Artificial Intelligence**]: Robotics—*Commercial robots and applications*

Keywords

children; domestic robots; field study; human-robot interaction; interaction design; wizard-of-oz experiment

1. THE "ROBJECTS" APPROACH

Designing domestic robots that can enhance daily lives of humans has been a long standing challenge in robotics and still it is not really clear what robots should and could do in the home [9]. User-centered research advocates that domestic robots need to be useful, easy to use, appealing, affordable, energy saving, and compatible with the home environment. Likewise, to ensure long-term adoption, a robot needs to meet user needs and prove its practicality within the "ecosystem" of a home. It is challenging to fulfill all criteria at once. Instead of trying to build novel multi-purpose robots, our approach is to design robotic objects, which we call *"RObjects"*. They integrate robotic functionalities into objects that are already part of our daily lives, and by this are expected to enhance the acceptance of robotic devices in homes. The overall aim of RObjects is to support users in specific daily routines in the domestic environment [11], and create meaningful human-robot interaction (HRI).

In our current project, we focus on robots for daily life in families with children. In a previous study [2], we saw that children generally respond very positively to robots and enjoy interacting with them, which has also been described by others [6, 7, 10, 13]. But where and how could a robot be an added value in a family home? In the initial phase of this project, we explored possible applications for RObjects in homes and identified that tidying up the children's toys is a challenge in a lot of families. We decided to create a RObject out of a toy box in which children would enjoy putting their toys back after playing (Fig. 1). We envisioned that, instead

HRI'14, March 3–6, 2014, Bielefeld, Germany.

http://dx.doi.org/10.1145/2559636.2559659.

of having a robot carrying out the tidying task for them, it could rather serve as an educative tool and motivate / help children to tidy up themselves. To bring children to tidy up would already be a big help in a lot of families, and could have a positive effect on the family ecosystem.

2. THE ROBOTIC TOY BOX "RANGER"

We call our prototype "Ranger". In French, the verb *"ranger"* means "to tidy up / sort out things" which suits the application of our robotic toy box.

2.1 Design Rationale and Related Work

A robot with an application in a family home must be designed for the entire family [9]. Even if Ranger's main users will be young children, we also need to consider parents, as both might perceive robots differently [14, 15]. Consequently, the robot's design, functionality and interaction should not only be effective and enjoyable for young children but also acceptable for parents. There were several issues to solve, both from a technical and human social side.

One fundamental question was what features Ranger would need to have and whether it should search and load toys itself? On one hand, in a home organization task (such as tidying), a robot is expected to automate the task to disburden humans. On the other hand, it should at the same time still allow people to feel like they have the control [9]. For this, the robot could rely on human-participation. In our tidying scenario the child should therefore remain the most active player, and consequently we did not equip the box with arms or hands that could grasp and load objects.

Another aspect was how Ranger should initiate an interaction and convey its intention to serve as a storage box for toys. To solve this, we can draw on a similar development, the "Child-Dependent Sociable Trash Box (STB)" [16]. This robot, which aims to engage children to collect trash and put it into it, was tested with 108 children (4-11 years old). The authors found three factors that were more effective for interaction and in conveying the STB's intention toward collecting trash: (1) the STB's ability to move (compared to an immobile trash box); (2) the boxes moving toward the trash instead of moving toward the children; and (3) a small group of STB's interacting with the children (compared to one single box). Regarding (3), we could not develop more than one prototype at this time, as we first wanted to find out how well our approach works in a real-world setting. Regarding (1), we equipped Ranger with two motorized wheels to make it mobile, so it could help transferring toys from one place to the other. Regarding (2), we thought about letting the box autonomously move toward a toy, in order to convey to the child to put the toy inside. The trade off of such an autonomous behavior is increasing complexity of the development and cost of the robot. Ranger would need vision and a larger on-board computational power, as well as obstacle avoidance strategies. Is it really necessary to have the robot being *proactive* in searching for and moving toward a toy itself? Could not just the child do this? These considerations led us to our main research question: *How does the robot's "pro-activeness" impact the interaction with the child and the motivation to tidy up?* To investigate this aspect in our case study, we modeled two different robot behaviors: a *proactive* robot that moves toward a toy and a *reactive* robot that responds to actions but more or less remains at one spot in the room.

Further, how should Ranger give feedback when the children put a toy? A similar but immobile toy storage box, called "Bubu Monstry" [8], which aims to encourage children (2-4 years) to return their toys back to it, makes chewing and eating sounds as soon as a toy is put. When all toys are put back, Bubu vibrates and burps. We decided to rather use sound and light cues to give positive feedback. In a study with Roomba [2], we saw that children liked the limited sound cues of the robot. We used interaction design methods to define a set of feedback cues for Ranger.

To let the robot display basic facial expression, movable eyes and eyebrows were added (Fig. 1). The eyes do not only give the robot a front but also a cartoon-like "fac" which can have a positive effect on the interaction [1]. We are aware that the presence of eyes on Ranger might also increase people's expectations. For this reason we avoided a more human-like face.

Finally, for this first design-evaluation iteration, we did not aim at developing a fully autonomous prototype but decided to use a human *"Wizard"* who would operate the robot from the background, following a Wizard of Oz method (Sec. 3.1).

2.2 Pro- and Reactive Systems

One concrete question we explored in this case study is whether a proactive robot triggers more and different kinds of interactions with children than a reactive robot (and in turn is more effective in the given task). The aspect of whether the system or the user takes the initiative is a relevant topic in the design of interactive systems. With being proactive we understand a system that tries to initiate an interaction (system driven) while in a reactive system, the user would take the initiative (user driven). The aspect of initiative- or turn-taking, dialogue, and how to balance the control between user and system is critical and a well studied topic in Human-Computer Interaction, Learning Technologies, Persuasive Technologies, and Interface Design. Ju and Leifer suggest that the degree of proactivity in a system and the required attentional demand from the user side should be adapted to the user's capabilities and need of control, to the situation, and to securtiy factors [5]. One way to balance the interaction in HCI is the implementation of so-called mixed-initiative. In HRI there is yet no common trend how to handle initiative taking, and most research on this deals with how to solve the issue from a technical point of view. Regarding our tidying scenario, the challenge is to find a good balance of the robot proactively trying to trigger interaction while still not pushing the children too much. In human-robot collaboration, it seems that robotic systems that use shared / collaborative control, similar to user-adapted autonomy, can positively impact the performance and the robot's usability [3]. However, there is no clear answer on how active or passive a robot should behave and certainly more research needs to be carried out that investigates the level of activeness in a robot which is beneficial to the human user and effective for the given task.

2.3 The First Prototype of Ranger

The first prototype of Ranger (Fig. 1) is a wheeled box (27 x 37 x 37 cm) with partial wooden surface. Being remote controlled, it can move around a flat surface, move its eyes and eyebrows, display colors (LEDs) and light patterns, and produce a limited set of sounds (integrated speakers).

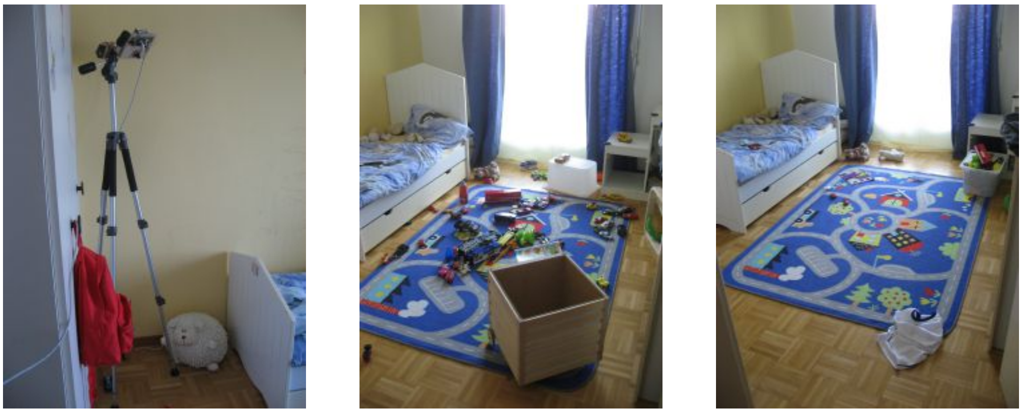

Figure 2: Study setup in one of the children's rooms; left: camera and embedded computer on a tripod are in one corner of the room; center: before participants entered the room, Ranger was next to the toys on the floor; right: the room after children had tidied up with Ranger and the study was over

During the field study in the children's rooms, Ranger was remote controlled by a human *"Wizard"* in a different room, using a video game controller connected to a laptop computer. The control program sent commands over Bluetooth to the robot's micro-controller. An external computer including a remote camera was placed in the children's room (Fig. 2), and connected to the laptop computer through Ethernet, so that the Wizard could observe the scenario. Fortunately, all the children's rooms were rather small, so that the robot was never out of the camera view and was thus always visible to the Wizard, who beforehand placed and adjusted the camera in the best possible way.

The Wizard followed a pre-defined script to model Ranger's feedback. These low-level robot behaviors were mostly pre-coded, and the Wizard pressed on a specific button of the remote control to execute them. This ensured that Ranger always reacted the same way to the same inputs:

- *Object put in box:* "rewarding" sound and red light pattern, occasionally wiggle-like move in parallel
- *Several objects after each other put in box:* "rewarding" sound, rainbow light pattern, box turns around itself
- *Object removed from box:* "emptying" sound and green light pattern
- *Box is touched or petted:* "blush" with pink light around the area of the "cheeks"
- *Box is kicked or mistreated:* "disturbance" sound and box moves backward
- *Box "finds" a toy on the floor* (only in *proactive* behavior): eyes rotate toward the toy, yellow pulsating light pattern, then after some seconds wiggle-like move, then back-and-forth move

The activity of removing a toy from the box should not be considered as "anti-tidying" because children would probably remove the toys from the box once it was full in order to store them in the drawer or shelf. However, we cannot say if the box was full based on the number of toys because of the different sizes of the objects.

3. METHOD AND PROCEDURE

3.1 Wizard of Oz Approach

We wanted to evaluate an early version of our robot prototype and study human-robot interaction in a realistic setting before developing a fully autonomous robot. To do so, Weiss *et al.* [12] propose the Wizard of Oz approach, in which the robot is not fully autonomous but controlled by a human operator who hides in the background.

Parents (but not children) were told beforehand that the robot was remote controlled. As we are mainly studying children's interaction with the robot, we do not feel that this biased the study. Rather, parents felt assured to have transparency about the functioning of the robot and knew it was safe. During the study, very few children noticed the tripod with the embedded computer and the camera on top of it (Fig. 2). As this "thing" did not resemble a usual camera, children did not notice that they were video-taped. Only two children asked what the "thing" was for and the experimenter replied that Ranger would need this "thing" to work properly. In general, it was the robot and not the tripod with camera that got children's attention, so we do not have the impression that the presence of the camera influenced children's behavior.

3.2 Study Design

We wanted to investigate not only how children use and interact with the robot in general but also whether and how variances in the robot's behavior impact the interaction, and how effective the tidying task was completed. Would a robot that proactively tries to trigger an interaction better motivate children to tidy up their toys, in comparison to a robot that shows rather passive, reactive behavior? To study these aspects, we configured two different robot behaviors:

- *Proactive:* Ranger tries to initiate interaction by driving around the room, searching for toys on the floor. When in front of a toy, it moves its eyes toward it and pulsates in yellow light. After about 5 seconds, if the child does not put the toy inside, Ranger also shows wiggle-like moves, and after 5 more seconds of no toy put inside, it moves a bit back-and-forth, with yellow pulsating light. By this behavior, Ranger proactively tries to trigger the action of putting a toy.

- *Reactive:* Ranger remains at one spot in the room (somewhere close to the toys on the floor) and reacts to the children's actions. The box hardly moves, does neither search for toys itself nor move toward children, but waits until a toy is put, and reacts to it as described above. By this, Ranger only shows a reactive behavior and does not try to initiate an interaction.

In one half of the families Ranger was operated in the *proactive* mode and in the other half in the *reactive* mode. This between-group design allows us to explore how the robot's different level of "activity" impacts the interaction and effectiveness of tidying up. In line with previous findings [16], we imagined that a more active system would trigger more interactions and better encourage children to tidy their toys. We formulated the following research hypothesis:

The proactive robot behavior will be (1) more engaging for children and (2) better motivate them to tidy up their toys (to put and remove objects).

As our analysis includes counting how many toys children put into and remove from the robotic box, we assumed a fairly comparable accumulated amount of toys among the two groups of each 7 families.

Participants

Besides investigating the impact of the robot's activity on task and interaction, we explored children's reaction to the robot and assessed parent's feedback to the design and acceptance of the robot. We tested Ranger in 14 families (1-4 children per family; in total 31 children, 17 parents). Between the two case study groups we balanced the number of children, age, and gender as much as possible. There were 16 boys and 15 girls from 2-10 years ($M = 5.3, SD = 1.98$). We are aware that the large age range implies different developmental stages in children. For this first evaluation of the robot we needed this diversity to explore at which age the robot's features are most effective. However, the large age range also makes it difficult to analyze and interpret data. Therefore we excluded age as a factor in statistical analysis but investigated age differences qualitatively.

Further, we had 16 parents (thereof 12 mothers, 4 fathers, mean age 40.5 years) with whom we conducted post-interviews (together with the children), and who filled in a short questionnaire to evaluate how they experienced and perceived the robot. Through this, we wanted to capture people's first impression of Ranger. Of course, a long-term study is required to study how far people's perception of the robot evolves over time. Still this spontaneous feedback is valuable and can help us to refine the design of this first prototype. The present family members were also asked to name specific aspects they liked and disliked about tidying up, about the robot, and how Ranger could be improved.

3.3 Course of the Study

Two researchers visited each family at their home for about one hour in total:

- *Introduction (~5 min):* Both researchers and the available family members introduced each other in the living room / kitchen. Parents were asked to sign a consent form (agreement that scenario was video-recorded, anonymized, and used for scientific purpose).
- *Pre-interview and preparation of setup (~15 min):* One researcher conducted a short interview with the family about the routine and challenges of tidying up, while the other researcher set up the robot and video recording in the children's room, as well as a remote control station in a different room.
- *Interaction study (max. 30 min):* The family went to the children's room to tidy up the toys with Ranger. No instructions were given on how to use the robot, the only goal was to tidy up the room. Children and parents were free to choose how to interact with Ranger and use it. When more or less all toys on the floor were tidied, the scenario was finished and participants left the room. Otherwise, the scenario was stopped after a maximum of 30 minutes. (Our focus was short-term interaction.)
- *Post-interview and evaluation (~5 min):* Another short interview and evaluation of the interaction and experience with Ranger concluded the visit. Parents evaluated various aspects of the robot and their experience on rating scales, and gave feedback on the design of the robot. The other researcher dismounted the setup.

To thank them for their participation, each child received a little gift at the end of the study before the researchers left the family's place.

3.4 Measurements, Coding and Data Analysis

All tidying scenarios with the robot were video-recorded and interactions analyzed by coding the videos according to a scheme that we developed. In contrast to how Kahn *et al.* [6] segmented behavior, we coded each behavior as one action as long as it continued. If several actions occurred several times within one minute, we coded several actions. Each of the children's actions towards the robot was categorized as one of the following: *exploration* (observing Ranger, trying to find out how it works), *misusage* (hitting or kicking it, poking it in the eye, trying to climb on it), *putting toy*, *removing toy*, using a *gesture* toward Ranger, *touching* it or *playing* with it (e.g. showing a toy, petting). In order to assess the quality of our coding scheme and to measure the reliability of our data, a second coder carried out a second coding pass on one of the 14 videos. We extracted and organized both coders' annotations in a spreadsheet and analyzed agreement on action-category as well as starting-time of the annotation. For this subset a Cohen's Kappa value of $\kappa = 0.74$ was obtained, which indicates a substantial agreement between the two coders. Children's verbal statements were not coded but analyzed qualitatively.

Interviews were audio-recorded for qualitative analysis. Adult participants rated their impression of Ranger in an 11-items questionnaire (Table 1). Most statements are adapted from Heerink *et al.*'s proposed questionnaire to measure the acceptance of assistive social robots [4]. We measured the following constructs: perceived ease of use (PEOU), perceived usefulness (PU), anxiety (ANX), attitude (ATT), intention to use (ITU), perceived enjoyment (PENJ), perceived sociability (PS), social presence (SP), and people's overall impression of the robot. To also get feedback on concrete aspects of the design of Ranger, both children and parents were asked to name 3 aspects they liked and disliked most about the robot, and to make suggestions for improving the robot.

Table 1: Questionnaire statements to evaluate Ranger (adapted from [4]), participants rated agreement on 5-point scale (1=agree, 5=disagree), mean M and standard deviation SD for each group, n=16 (8 per group)

	construct	statement	"proactive" M (SD)	"reactive" M (SD)
1.	PEOU	I think Ranger is easy to use.	1.88 (0.83)	**1.00** (0.00)
2.	PU	Ranger is useful.	1.63 (1.06)	**1.50** (1.07)
3.	PU	It would be convenient for me to have the robot.	**1.75** (1.04)	2.00 (1.20)
4.	ANX	I would be afraid to make mistakes with Ranger or to break something.	4.50 (1.07)	**4.29** (1.11)
5.	ATT	I think it's a good idea to use Ranger.	**1.25** (0.46)	1.75 (1.16)
6.	ITU	I think I would use Ranger during the next few days.	**1.25** (0.71)	1.88 (1.13)
7.	PENJ	I think Ranger is boring.	4.88 (0.35)	**4.50** (1.07)
8.	PENJ	I enjoy when Ranger is responding to me.	1.13 (0.35)	**1.00** (0.00)
9.	PS	I think Ranger is nice and pleasant to interact with.	1.50 (1.07)	**1.38** (0.52)
10.	SP	I can imagine Ranger to be a living creature.	2.75 (1.49)	2.75 (1.49)
11.	other	My overall impression of Ranger is: (1 = very good, 5 = very bad)	1.63 (0.74)	**1.38** (0.52)

4. RESULTS

By coding the 14 videos with a total length of about 3 hours, 1740 distinct actions were obtained. The scenario duration varied between 5-27 min ($M = 704\ s, SD = 420\ s$) with no significant difference between the two groups (*proactive* $M = 778\ s$, $SD = 425\ s$; *reactive* $M = 625\ s$, $SD = 415\ s$); ($t(29) = -1.01, p = .32$; independent samples t-test). Thus, overall children did not tidy up quicker in one of the two groups.

4.1 Children's Interaction with the Robot

How Children Approach and Make Sense of Ranger

Children naturally understood how to use Ranger, regardless whether it tried to initiate an interaction itself (*proactive*) or not (*reactive*). They started putting toys into the box on average after 2:22 min ($SD = 245\ s$), with no significant difference with robot behavior ($t(13.4) = -1.5, p = .16$; independent samples t-test, unequal variances). There was the exceptional case of a 4 years old girl who was very shy and hesitated putting something into the robotic box for 27 min. Asked why, she explained that none of her toys would belong into this box but only to a specific place on the shelf.

It seemed that *initially* it was more obvious for the children to put a toy in the proactive behaving robot, as it was actively moving toward the toys. Before putting a toy, most children first shortly explored the robot by looking inside or underneath it, or they touched it to investigate the reaction.

There was a great enthusiasm for the box, and only the youngest children (up to 2 years) were afraid when the robot started to move or wiggle. We observed this reaction of very young children also in our previous work with vacuum-cleaning robots [2]. It seemed that the children up to 2 years could not immediately make sense of the robot. However, after watching how the older brother or sister interacted with the robot, they also approached it, and eventually also put a toy. In general, there were different strategies of making sense of Ranger, as the *proactive* robot enabled more varied interaction and allowed transportation of the toys, for instance. The proactively moving robot helped children in starting and structuring the process of tidying up, which seems to be a major difficulty (*"Where and how do you start sorting out things when there is a huge mess?"*).

In the following, we present the interaction dynamics for one case from each robot behavior. The two examples are not representative for the respective group but were chosen as they are fairly comparable (number / gender of kids; duration of scenario).

Family 1 (boy B (5), girl G (3), *proactive robot*
Ranger moves in front of toys on the floor, both children are attracted and explore the box; B touches Ranger's eyes, puts a toy (after 1 min); B and G are very fascinated by the robot's feedback and start putting a lot of toys in a series (22 toys in 3 min); G watches; there is another toy left on the floor, Ranger moves toward it; this proactive robot behavior motivates B to play a game: B removes a toy from the box and places it somewhere on the floor to make Ranger go there; each time, B jumps happily when Ranger finds the toy on the floor; B eventually puts a toy from box in drawer but still distributes toys on the floor to make the robot go there; [comment: this is playful but not tidying up]; total 50 toys put/removed in 13 min

Family 2 (boy B (8), girl G (3), *reactive robot*
B and G look at Ranger; B waves in front of the robot which doesn't react; B puts a toy (after 1 min,); both kids are fascinated by Ranger's feedback and put a lot of toys (57 toys in 5 min); the box is reasonably full and B starts removing them; also G removes toys and puts them back on the floor; B actually tidies up the toys that G puts on the floor; G continues removing (56 toys in 5 min) and B tidies up the things into drawers and other boxes; B explores the robot; B and G continue until the box is fairly empty again; [comment: kids were focused on tidying but were disappointed that the box didn't move around]; total 122 toys put / removed in 11 min item

Figure 3: Children interacting with Ranger during the field study: The robot's eyes received remarkable attention. Left: first moments of a family exploring together the robot; center: two boys putting toys into Ranger, which displays red lights; right: a girl showing a toy to the robot

How Children Explore and Interact with Ranger

Concerning the total duration of all coded actions, almost half of the time (47 %) children explored the box. Though the differences were not significant, the *proactive* robot was explored almost twice as much as the *reactive* robot. One of the most prevalent interactions was the exploration of Ranger's eyes. Several children touched the robot's eyes or "showed" toys to Ranger before putting it (see Fig. 3). Others waved their hand in front of the robot's face, to find out how it would react and whether the robot would be able to see. With the *proactive* robot, children were fascinated by seeing the box moving around their room and some invented games to let it drive around (e.g. putting a toy on the floor, so that the robot would "search" for it). It seems that children perceived the proactive robot as more "companion-like" and probably this engaged some children in socializing with it. Though we have no real evidence for how "social" children perceived the robot, two boys waved at Ranger and said "bye-bye" when they were leaving the room after the scenario.

We also qualitatively investigated gender and age differences in how children interacted with the robot. It seems that younger children (3-6 years) benefit most from using Ranger. They were more fascinated by the audio and light cues than older children (7-10 years). Children less than 3 years old did not really use the box for tidying but rather watched it or held on it to practice walking around. There were interesting qualitative gender differences. Boys more often mistreated and gestured toward Ranger than girls, who slightly more often petted the robot. For instance, several boys wanted to direct the box around using pointing gestures and verbal comments, such as *"go over there!"* or *"come here"*. Nevertheless, each child has her/his own character and we cannot generalize these findings. However, again, we made similar observations in our previous work with children interacting with vacuum-cleaning robots.

4.2 Tidying up the Toys with Ranger

Concerning the average number of toys put into / removed from the proactive and reactive robot (Fig. 4), there was in general more activity within the first couple of minutes. This could be related to the *novelty effect*. On average children put / removed more toys in the *reactive* robot; even toward the end. An explanation for the rather low activity of putting / removing toys in the *proactive* robot is that children were more playful with the moving robot and seemed less interested in tidying.

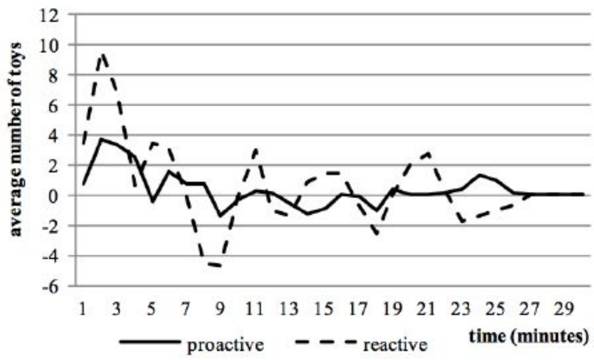

Figure 4: Average number of toys put into and removed from the robotic box per minute; more fluctuations with the reactive robot; in both cases the activities become less over time

Regarding the total number of occurrences of the coded actions, 48 % were "put" and "remove" actions. This generally supports our approach that Ranger can be used to motivate children to tidy up their toys, at least in short term.

Contrary to part (2) of our research hypothesis, data suggests that it was the reactive and not the proactive robot behavior that better motivated children to tidy up: Marginally significantly more toys were put and removed in the *reactive* compared to the *proactive* robot (compared means using ANOVA) (see Table 2). The respective Cohen's d-values indicate moderate effect sizes. Apart from putting or removing toys, all the other actions that were not directly related to tidying up (explore, misuse, touch / play, gestures) were carried out more often with the proactive robot. These differences were not significant, though for "misuse" and "gesture to robot" moderate effect sizes were found. Overall, this result is in line with part (1) of our research hypothesis: The proactive robot seems to be more engaging for children.

More generally, our data suggests that both a proactive and a reactive robot behavior can trigger interactions, however different kinds of interactions. Interestingly, it seems that the *reactive* robot was more effective in enhancing tidying behavior, whereas the *proactive* robot motivated children far more to explore it, misuse it, touch and play with it, and

Table 2: Average number of actions per child carried out with the proactive (n=16) and reactive (n=15) robot; task-related actions (put/remove toy) are carried out more often with the reactive robot, while playful actions (explore, misuse, touch and play, gesture to robot) are more often carried out with the proactive robot; statistical analysis in bottom part; data obtained by coding children's actions from the video recordings

robot behavior	**put toy** M (SD)	**remove toy** M (SD)	**explore robot** M (SD)	**misuse robot** M (SD)	**touch and play** M (SD)	**gesture to robot** M (SD)
"proactive"	9.81 (10.13)	4.94 (6.12)	**18.38** (15.18)	**4.50** (8.08)	**9.81** (11.19)	**2.81** (4.10)
"reactive"	**24.80** (27.41)	**16.13** (23.18)	12.80 (6.33)	0.80 (1.15)	6.80 (5.77)	1.07 (1.28)
F(1,29)	4.18	3.48	1.74	3.09	.869	2.49
p-value	**.05**	.072	.2	.09	.36	.126
Cohen's d	**.73**	**.66**	.48	**.64**	.34	.57

use gestures toward it (Table 2). We interpret that a *proactive* system is more fun for children but might distract from the main task (here: tidying). For this, a *reactive* systems seems to be more efficient.

4.3 Evaluation: Acceptance and Design

In the subsequent interview, both parents and children expressed that they liked Ranger and that they enjoyed interacting with it. Parents appreciated that the robot engaged the children in a positive and meaningful way; they liked our general idea of making a RObject out of a toy box, such that it would make tidying up a playful activity.

Overall, parent's ratings (see Table 1) show that Ranger is perceived as very acceptable: there is a general agreement that Ranger is easy to use, useful, and that it would be convenient to have the robot. Parents could also imagine using the robot during the next few days, and they would not be afraid of making mistakes with it or breaking something. The interaction was evaluated as enjoyable, and not boring. Interestingly, participants were undecided about whether Ranger could be considered as a living creature.[1] These ratings did not differ significantly between the two groups (independent-samples t-test). Only for "perceived ease of use", a significant difference was found ($t(7) = 2.97, p = .021$, unequal variances, due to the fact that all participants provided the very same rating (1.0) for the reactive robot) (see Table 1). Surprisingly, participants rated the *reactive* robot as slightly easier to use than the *proactive* one. However, given the limited number of participants who filled in the questionnaire (n=16, 8 in each group), these are not strong results.

There were very few negative remarks. With the *reactive* robot, participants disliked that the robot was hardly moving. One father also criticized that our prototype was very limited because *"it doesn't do things on its own"*. He admitted that he had much higher expectations of "a robot". Contrary, what participants liked most about Ranger were its light and sound cues, the moving eyes, and its simple design (the fact that it was *"just a wooden box"*). This supports our general approach of RObjects. To improve Ranger, participants suggested to include more sounds, and also use verbal statements. Further, more than one box would be useful to sort toys better.

During the conversation, children attributed emotional states to the robot, describing it would be "happy" when they put a toy and "not happy" when it was empty. This shows that yet the limited audio and visual cues were appropriate and perceived as being emotional. It suggests that children viewed the robot as a kind of social agent with mental states. However, the attributions could also be due to a reflection of the child's own feelings and experience.

[1]This "neutral" result shows that it is not easy to evaluate something as abstract as the "life-likeness" of a robot. It doesn't seem appropriate to keep asking whether a robot is either "alive" or "not alive" because it might appear to be somehow both. See [6].

5. DISCUSSION AND CONCLUSION

We presented the design and evaluation of our robotic toy box prototype "Ranger" that aims to motivate children to tidy up their toys.

Our study is limited by a rather small sample size (which is also critical for the statistical analysis), quite big age differences in the children, and the fact that only one researcher coded the videos and interactions with the robot. Further, we assumed that there was a fairly equal number of toys across the families in the two study groups.

Even though Ranger was able to motivate children to tidy up, parents critically remarked that putting everything in one box or removing toys to put them back on the floor is not tidying. Also, the positive effect might be due to a *novelty effect* during this first interaction. Fascination might fade out quickly and we don't know yet how the interaction changes over time. In the future, long-term studies are needed to investigate how far our results can hold over time.

Several people were surprised that though we consider Ranger a "robot", it does not carry out tasks on its own. However, we do not aim at having a robot carry out a task *for* the family but *with* them. Especially with children, who are developing plenty of different skills and taking responsibility when they tidy up, we believe that it is important to let them do their job themselves, encouraged by a little robot. Still, deploying a robot in a children's room raises ethical questions, in terms of family roles and especially as young children seem to readily engage with robots in a social way.

Findings of our study suggest that both a *proactive* and a *reactive* robot can be engaging for children, however, in a different way. The *proactive* robot triggered more playful behavior but also seemed to distract from the tidying task. In contrast, the *reactive* robot triggered more tidying behavior but engaged children less in playing with it. Consequently, we think that probably a robot that tries to balance initiative-taking with the user could result in an engaging but still purposeful behavior (e.g. serious game).

Two further aspects of this project which we are currently investigating are (1) how various kinds of robot behavior (e.g. unexpected robot behavior) impact child-robot interaction, and children's perception of cognitive abilities of the robot, with the goal of sustaining interaction over a longer period of time, and (2) the development of an autonomous version of Ranger which features shared control, such that it is possible to conduct a long-term study with Ranger in family homes.

Acknowledgements

We thank all the families who participated in this study. We also thank Maya Cakmak and the reviewers for their help in refining this article, as well as Frédéric Kaplan and Fanny Riedo for their ideas.

This research was supported by the Swiss National Science Foundation through the National Centre of Competence in Research Robotics.

References

[1] J. Fink. Anthropomorphism and human likeness in the design of robots and human-robot interaction. In S. S. Ge, *et. al.*, editors, *Social Robotics*, no. 7621 in Lecture Notes in Computer Science, pp 199–208. Springer, Jan. 2012.

[2] J. Fink, V. Bauwens, F. Kaplan, and P. Dillenbourg. Living with a vacuum cleaning robot. *International Journal of Social Robotics*, 5(3):389–408, 2013.

[3] T. Fong, C. Thorpe, and C. Baur. Collaboration, dialogue, human-robot interaction. In P. R. A. Jarvis and P. A. Zelinsky, editors, *Robotics Research*, no. 6 in Springer Tracts in Advanced Robotics, pp 255–266. Springer Berlin Heidelberg, Jan. 2003.

[4] M. Heerink, B. Kröse, V. Evers, and B. Wielinga. Measuring acceptance of an assistive social robot: a suggested toolkit. *RO-MAN 2009*, pp 528–533. IEEE, Sept. 2009.

[5] W. Ju and L. Leifer. The design of implicit interactions: Making interactive systems less obnoxious. *Design Issues*, 24(3):72–84, 2008.

[6] P. H. Kahn, Jr., B. Friedman, D. R. Pérez-Granados, and N. G. Freier. Robotic pets in the lives of preschool children. *Interaction Studies*, 7(3):405–436, Jan. 2006.

[7] T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro. Interactive robots as social partners and peer tutors for children: a field trial. *Hum.-Comput. Interact.*, 19(1):61–84, June 2004.

[8] The Interaction Lab - Holon Institute of Technology. Bubu Monstry, URL: http://interaction.hit.ac.il./bubu-monstry/.

[9] C. Pantofaru, L. Takayama, T. Foote, and B. Soto. Exploring the role of robots in home organization. *HRI 2012*, pp 327–334, New York, NY, USA, 2012. ACM.

[10] B. Robins, K. Dautenhahn, C. L. Nehaniv, N. A. Mirza, D. François, and L. Olsson. Sustaining interaction dynamics and engagement in dyadic child-robot interaction kinesics: lessons learnt from an exploratory study. *ROMAN 2005*, pp 716– 722. IEEE, Aug. 2005.

[11] F. Vaussard, M. Bonani, P. Rétornaz, A. Martinoli, and F. Mondada. Towards autonomous energy-wise RObjects. In D. Hutchison, *et. al.*, editors, *Towards Autonomous Robotic Systems*, vol. 6856, pp 311–322. Springer Berlin Heidelberg, 2011.

[12] A. Weiss, R. Bernhaupt, D. Schwaiger, M. Altmaninger, R. Buchner, and M. Tscheligi. User experience evaluation with a wizard of oz approach: Technical and methodological considerations. *Humanoids 2009*, pp 303–308. IEEE, Dec. 2009.

[13] A. Weiss, D. Wurhofer, and M. Tscheligi. "I love this Dog"—Children's emotional attachment to the robotic dog AIBO. *International Journal of Social Robotics*, 1(3):243–248, June 2009.

[14] S. Woods and K. Dautenhahn. Child and adults' perspective on robot appearance. *AISB 2005*, pp 126–132, University of Hertfordshire, Hatfield, England, 2005. SSAISB.

[15] S. Woods, K. Dautenhahn, and J. Schulz. The design space of robots: investigating children's views. *ROMAN 2004*, pp 47– 52. IEEE, Sept. 2004.

[16] Y. Yamaji, M. Taisuke, Y. Yuta, R. De Silva, and M. Okada. STB: child-dependent sociable trash box. *International Journal of Social Robotics*, 3:359–370, Oct. 2011.