

A Baby Monitor as a Shape-Changing Interface

Simon SCOTT-HARDEN
School Of Design, University Of Northumbria
Newcastle Upon Tyne, UK
Simon@Zelo.Co.Uk

Ben SALEM
School Of Engineering, University Of Liverpool
Liverpool, UK
Mail@Bsalem.Info

In this paper we present a comparative study between a commercial Baby Monitor (model BT 250), and a baby monitor designed and built as a Shape-Changing Interface. We explain the study and present its results, followed by a discussion and some conclusions regarding our interaction design approach.

Shape-Changing Interface, Tangible User Interfaces, Metaphor.

1. INTRODUCTION

Shape-Changing Interfaces (SCI) were first defined as TUI with changing appearance to reflect internal states (Rasmussen et al., 2012), they connect back and forth the digital and the physical (Poupyrev et al, 2007). We see Shape-Changing Interface (SCI) as a type of Tangible User Interfaces (TUI), and as physical objects with an aesthetically sophisticated and pleasing form. We believe that the SCI design should also take into account that the form will evoke an affective response. Harmony, unity and agreement between the elements of the form should be sought after, bearing in mind that usage and possession of beautiful design provide sensory pleasure and stimulation (Girouard et a., 2013). Our design approach aims to deliver:

- A TUI with changes of physical attributes such as shape, appearance and temperature, merging inputs and outputs
- A Metaphoric representation of content matching some of the users' intentions, and implementing interaction tasks that are a direct translation of these intentions
- Having a Language-like user's inputs based on common actions, gestures and manipulations
- Being engaged in 'Form-language' outputs, without human agency, to render content and interaction clues

2. BABY MONITORS

A baby monitor is generally described as a two parts device that helps you to keep supervising a baby while in a different room, through sound,

lights and information displayed on an LCD screen. It also monitors aspects of the baby's environment (e.g., room temperature) and lets the user remotely play a lullaby and talk to the baby.

Table 1 lists the key functions of the BT250 Baby Monitor and their SCI equivalents. The equivalence here was determined by looking at the functions of baby monitors as reported in a survey of 14 first-time mothers and by determining what were the essential features to be implemented in our device.

Table 1: Baby Monitor and SCI functions

Functions	Devices	
	Baby Monitor	SCI
Temperature Display	LCD display	Base's LED glow & colour
Set Temperature	Not applicable (via thermostat)	Hold petals between hands to warm up and blow on device to cool down
Lullaby	Navigate through menu	Press Stamen (torch)
Speech	Press & hold side button	Press & hold stamen (torch)
Noise Level	LED display	Moving (breathing) and glow of petals
Torch	Press button on rear of device	Stamen is a torch to be taken away

Our survey results indicated that sound channel and volume control (50% of respondents), playing lullaby (42%), room temperature (35%), torch (35%) and two-ways talk (14%) were the most sought after features. Since the BT 250 baby monitor did not include a mean to control the room temperature, we have added a thermostat next to

the Baby Monitor that is used to set the temperature in the baby's room.

2.2 Design of our Baby Monitor as a SCI

Having selected the functions of the baby monitor, we then designed a new SCI accordingly.

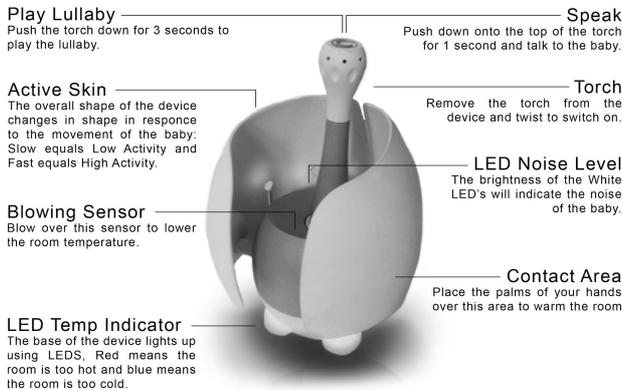


Figure 1: The SCI Baby Monitor

3. EXPERIMENTS

We set up our experiments in a Wizard of Oz style, with venue and scenario as close as possible to a real world babysitting case (albeit no babies were involved). Both the BT 250 and the SCI baby monitors were used in a living room, and the baby and the remote baby room were simulated. We conducted empirical testing within the experiment venue and we collected both objective measurements in addition to the subjective data obtained from questionnaires and interviews.

3.1. Test conditions

We have defined four test conditions, which we refer to as Normal, Medium, High and Highest. Table 3 lists the conditions and the corresponding environment features.

Table 3: The Four test conditions of the experiment

Conditions	Parameters		
	Temperature	Noise	Arousal
Normal	20 °C (normal)	None	Low
Medium	26 °C (hot)	Medium	Medium
High	14 °C (cold)	Medium	Medium
Highest	20 °C (normal)	High	Awake

Furthermore, for each of the test conditions, test subjects were given clear explanations of the scenario and accordingly action instructions (see Table 4).

Table 4: User's actions according to the test conditions

Conditions	Scenario	User's Actions
Normal	All is fine	None
Medium	Room is too hot, baby is slightly upset	Lower the room's temperature, play a lullaby to soothe the baby
High	Room is too cold, baby is half awake	Increase the room's temperature, talk to the baby
Highest	Baby is upset and fully awake	Take the torch, go to see the baby

Finally, test subjects were also given detailed instructions as to how they should use either of the Baby Monitors.

3.2 Method

The experiments were carried out in a test room - the lounge of the house, where participants were asked to watch the movie "Madagascar" (Directed by Darnell & McGrath, 2005), which started at the same scene for all the participants. Participants were told to behave as if there is a baby in one of the bedrooms upstairs and the remote part of the baby monitor was installed in that bedroom. The aim was to have the participants staying in the lounge in an as normal a setting as possible. For the whole duration of the experiment there was a video recording of the room, and participants were informed of this.

3.3 Test Subjects

We selected 23 participants aged 20-35, enrolled as either under-graduates or post-graduates at the authors' University; thus matching the profile of babysitters, in other words our target users, as encountered in the real world. All of the participants had no preconceived ideas or values of baby monitors prior to the study, apart from a passing knowledge of this type of device. They all however had no issues with babysitting, which was potentially an income generating activity they could adopt. Prior to the experiments, participants were asked the following question: "do you know how to operate a baby monitor?" (In this question, 2 is very much and -2 is not at all). Participants gave the following mean answer: -1.2609 (SD 0.9637) indicating that the majority of the participants did not know how to operate a baby monitor prior to the experiment. Furthermore, 65% of participants had babysitting experience prior to the experiment though not acquired regularly; and 87% of the participants have not received training in babysitting.

3.4. Experiment Procedures

Each participant underwent the same procedure for all the experiments. They took part in the study with

the commercial and the SCI baby monitors, going through the different test conditions. Both the devices (i.e. the experiment stage) and the test conditions (i.e. scenarios) were randomly ordered. Each test condition corresponding to one of the scenario combined with one of the devices lasted up to 90seconds. The test conditions started when the participants heard a continuous buzzer, and ended when they heard a two-beep buzzer. There was a 30 seconds break between the test conditions. At the end of the first stage, when all the scenarios were simulated for the device present in the room, participants were asked to move back to the briefing room. This is when the devices were swapped.

The second stage of the experiment ran in an identical manner, with this time with the other device present in the room. Afterwards, participants were given a final questionnaire. They were then asked if they had any further comments, before one of the experiment staff entered in a casual conversation with them discussing the devices used, the experiments and their general impressions and opinions. Subjects were then given free transportation to a destination of their choice, as well as a £10 shopping voucher.

3.5 Objective Measurements

For each of the test conditions, using either device, we have measured the following times:

In the normal scenario, where no action was required, no participant picked up either of the devices, and the test conditions lasted 90 seconds.

T1 – Time between system ready and simulating scenario and device is picked up by user. We measured the time from when the system was ready and when the participants picked up the device.

T2 – Time device is actually used (e.g. hand held, blown onto...). We measured the time from when the users hold the device, some interaction occurs, and when the device is put back on the table with a capping on at 90 seconds.

We then compared the means using Paired-Samples T test, we obtained the following results (standard deviations are in parentheses):

For **times T1**, there was no statistically significant difference between the BT 250 Baby Monitor (B/M) and our device (SCI).

For the “Medium Scenario” test condition, we found that the mean T1_{B/M} is 21.06 (12.44), the mean T1_{SCI} is 22.66 (13.50), and that the mean difference is 1.60 (N=15, Sig. P=0.685).

For the “High Scenario” test condition, we found that the mean T1_{B/M} is 15.54 (9.62), the mean T1_{SCI} is 17.68 (8.62), and that the mean difference is

2.13 (N=22, Sig. P=0.457).

For the “Highest Scenario” test condition, we found that the mean T1_{B/M} is 11.04 (5.41), the mean T1_{SCI} is 12.66 (5.48), and that the mean difference is 1.62 (N=21, Sig. P=0.277).

For **times T2**, we have found the following results:

For the “Medium Scenario” test condition, we found that the mean T2_{B/M} is 25.23 (11.02), the mean T2_{SCI} is 13.69 (4.83), and that the mean difference is -11.53 (10.03) **in favour of our device** and this is highly significant (N=13, Sig. P=0.001).

For the “High Scenario” test condition, we found that the mean T2_{B/M} is 21.47 (9.13), the mean T2_{SCI} is 13.63 (5.20), and that the mean difference is -7.84 (10.01) **in favour of our device** and this is highly significant (N=19, Sig. P=0.003).

For the “Highest Scenario” test condition, we found that the mean T2_{B/M} is 16.40 (3.68), the mean T2_{SCI} is 21.20 (11.95), and that the mean difference is not significant (N=15, Sig. P=0.107).

3.6 Subjective Measurements

We also ran a paired-samples T-test analysis on the questionnaires' answers. We have asked a series of questions, after using both devices. We have obtained the following results. In these questions, 2 is strongly disagree and -2 is strongly agree. Thereafter a negative answer is agreeing with the statement of the question, while a positive answer is disagreeing. We asked eight questions to the participants as follow:

Q1 “*I found the device unnecessarily complex*”.

Q2 “*I think that I would need the support of a technical person to be able to use this device*”.

Q3 “*I thought there was too much inconsistency in this device*”.

Q4 “*I would imagine that most people would learn to use this device very quickly*”.

Q5 “*I felt very confident using the device*”.

Q6 “*I need to learn a lot of things before I could get going with this device*”.

Q7 “*While using the device I felt I was manipulating information about the baby rather than manipulating a device*”.

Q8 “*I found that the device clearly reacted to my actions*”.

The participants gave us answers as reported in Table 5.

Table 5: Mean answers for questions 1-8

	Mean Answers		
	B/M	SCI	Mean Difference

Q1	1.4091	0.2727	1.1364 (N=22, t=-3.834, df=21, Sig. p= 0.001) B/M
Q2	1.6087	1.0435	-0.5652 (N=23, t=-2.260, df=22, Sig. p= 0.034) B/M
Q3	1.3182	0.5909	-0.7272 (N=22, t=-2.296, df=21, Sig. p= 0.032) B/M
Q4	-1.5217	-0.9130	0.6087 (N=23, t=2.440, df=22, Sig. p= 0.023) B/M
Q5	-0.9545	-0.0909	0.8636 (N=22, t=2.310, df=21, Sig. p= 0.031) B/M
Q6	1.2609	0.4348	-0.8261 (N=23, t=-2.702, df=22, Sig. p= 0.013) B/M
Q7	0.7500	-0.8500	-0.8500 (N=20, t=-2.669, df=19, Sig. P= 0.015) SCI
Q8	-1.5000	-0.8636	0.6363 (N=22, t=2.536, df=21, Sig. P= 0.019) B/M

When comparing the mood of the users as reported in bipolar-scaled questionnaires, we ran within subjects, one-way ANOVA using the Repeated Measures Analysis in the General Linear Model. We then analysed the results running a Mauchly's Test of Sphericity where we identified results that were either significant or not. We can report on the mood question "relaxed or tense": The Mean Scores (N=23) for the moods before experiments (1.0870), after using the Baby Monitor (0.4783), and after using the SCI (0.8696) differed significantly: $F(2,44) = 4.453$; $p=0.17$, Partial eta squared = 0.168, which is a large effect. This result is **in favour of the SCI**.

We also ran a comparative questionnaire asking what device would participants prefer (-2 for B/M and +2 for SCI, note that only device name separated by five boxes were shown to participants, not the -2 and +2 values) and we have collected the answers as shown in table 6.

Table 6: Device Preference

Which Device	Answers	
	Mean (Standard Deviation)	Winner
<i>I prefer</i>	0.13 (1.4239)	SCI
<i>Is best ergonomically designed</i>	0.38 (1.3219)	SCI
<i>I will recommend to a friend</i>	0 (1.2344)	None
<i>Is more reliable</i>	-0.31 (1.2105)	B/M
<i>Best fits in its surroundings</i>	0.56 (1.3759)	SCI
<i>Does not require too much effort to operate</i>	-0.08 (1.4432)	B/M

4. DISCUSSION

First, even without prior training, participants were able to use our device in comparable performance to a commercial baby monitor. In all the tests, the task completion ratio was not significantly different between devices. Second, in our objective measurements, the time taken using the devices is significantly in favour of our baby monitor in two of

the three test conditions. Anecdotal evidences suggest that in the case of the "Highest Scenario", the difficulty was that users were reluctant to pickup the SCI baby monitor stamen (torch) as instructed.

The participants' answers to Q7 indicate that we were successful in implementing the SCI as a metaphorical representation of the baby while asleep in a remote room. This study has confirmed the validity of our Industrial Design of the physical device, and the validity of the Interaction Design of the shape-changing interface. However, interestingly all the other answers that were statistically significant were in favour of the commercial baby monitor. Although, when the commercial and the SCI baby monitors usage times are compared, we see that the measurements are in favour of our prototype in two of the three scenarios. And when we compare task completion between the two devices, we see that there is no statistically significant difference in the number of participants that can complete the tasks with one device or the other. Furthermore, the only statistically significant difference in terms of mood is the "relaxed or tense" mood. Participants report that they are better relaxed after using the Active Form.

The participants' reporting seems to indicate that they perform well with both devices and also that they are more relaxed after using the SCI and feel more like "manipulating information about the baby rather than manipulating a device" (Q7). However, they report that the commercial baby monitor is better (Q1 to 6, and Q8). Our impression was reinforced in additional comments filled on the forms handed out, where we have noticed interesting comments, such as: "In general the [SCI] device was easier to use and looked better" and "I was much more aware of room temperature with the [new] device as opposite to the baby monitor".

5. REFERENCES

- Girouard, A., Vertegaal, R., Poupyrev, I. 2013. Special Issue: Organic User Interfaces, Interacting with Computers 25(2): 115-116.
- Poupyrev, I., Nashida, T., Okabe, M. 2007. Actuation and tangible user interfaces: The Vaucanson duck, robots, and shape displays. In Proceedings TEI, Baton Rouge, LA, USA, 15-17 February 2007, 205-212.
- Rasmussen, M.K., Pedersen, E.W., Petersen, M.G., Hornbaek, K. 2012. Shape-changing interfaces: a review of the design space and open research questions. In Proceedings CHI 2012, 735-744.