Physical Fidelity: Exploring the Importance of Physicality on Physical-Digital Conceptual Prototyping

Joanna Hare[1], Steve Gill[1], Gareth Loudon[1], Devina Ramduny-Ellis[2] and Alan Dix[2]

[1]The National Centre for Product Design and Development Research (PDR), The University of Wales Institutue, Cardiff (UWIC), Western Avenue, Cardiff, CF5 2YB, UK.
{juhare-pdr, sjgill, gloudon}@uwic.ac.uk.
[2] InfoLab21, Lancaster University, Lancaster, LA1 4WA, UK.
devina@comp.lancs.ac.uk, alan@hcibook.com.

Abstract. The physicality of digital-physical devices is an essential part of our interaction and understanding of information appliances. This paper draws on the findings of an empirical study investigating the effect of physical fidelity on a series of user trials. Three prototypes of a single design intent were built, the standard of their construction dictated by the time imposed on the designer. In choosing this constraint, the authors present the argument that the most important driver in decisions that dictate fidelity levels is the available and/or necessary time required for making a prototype in order to generate information of the right quality. This paper presents the empirical and qualitative results of the trials, which suggest that there is little effect of fidelity on user performance, but the user's ability to give constructive feedback on the design was influenced by the nature of the prototypes.

Keywords: Physicality, prototyping, fidelity, information appliance, product design, tangible interface, low fidelity prototyping.

1 Introduction

This paper focuses on *information appliances,* devices designed to do one task, but do it well. The design of these devices poses interesting challenges to the design community because not only do information appliances have physical considerations (size, shape, buttons, etc.), they also have digital considerations (dedicated computer running software menus, features, function etc.). The digital and the physical are therefore inescapably linked in information appliances.

Prototypes are used to physically explore an idea very early on in the design process and *interactive* prototypes can be used to explore the digital considerations integrated with the physical form. However, prototypes at this initial stage are inherently 'quick and dirty': they should not require a lot of time to make and should be an exploration of an idea rather than a refined model (what Schrage [1] describes as *'Serious Play'*).

There are many academic and industry research groups working on tools and techniques for rapid interactive prototyping. These include:

- *Paper Prototyping* [2] – a very low tech approach requiring no technical skills; the user usually interacts with a paper-based version of the interface on a physical model and the screen is adjusted by a facilitator, acting as the 'computer'.
- *D.tools* [3] – a toolkit with bespoke hardware and software.
- *Phidgets* [4] and *Arduino* [5] – both provide electronic 'building blocks' to integrate into a prototype.
- *IE (Information Ergonomics) System* [6] – a flexible system of hardware and software linking a prototype to a PC.

One of the underlying recognitions that tie all this work together is that prototypes need to be made quickly in order to evaluate the tangible interactions.

The *fidelity* of a prototype is usually considered to be the resolution (the refinement and detail) of the model. A number of publications have been focused on the effect of fidelity and the advantages and disadvantages of different prototyping techniques. Sefelin et al. [7] looked at the user's willingness to criticize paper prototypes versus their willingness to criticize computer based models. Virzi et al. [8] found that there was little difference in usability data for high and low fidelity models of standard two dimensional graphical interfaces and an interactive voice response system. McCurdy et al. [9] argued for a mixed approach that allowed various aspects of a prototype to be built at different fidelity levels according to the design component being prototyped. They go on to suggest that there are five 'dimensions' or fidelity aspects that can be defined as somewhere between high and low within the same prototype, namely, aesthetics, depth of functionality, breadth of functionality, richness of data and richness of interactivity. So far this concept of mixed fidelity has been trialed with software but not physical prototypes.

Information appliances and therefore prototypes of information appliances are inherently physical. Physicality as a term, is becoming more recognized with two International Workshops on Physicality [10, 11] held recently, plus Don Norman's article on Physicality [12]. Physicality is loosely understood as being the physical nature of something, for example, a form, process or button.

This paper seeks to contribute to our understanding of the nature of physicality in the design of information appliances so that designers can become more aware of when and how to use it. To this end, we explore physicality in the context of fidelity through user trials conducted on a conceptual information appliance.


2 Background

Gill et al. [13] conducted a number of trials on a wireless home phone. They demonstrated that low fidelity physical prototypes can produce similar usability results as the end product, thus significantly outperforming touch screen mock-ups. They went on to test prototypes of decreasing fidelity until they reached a point where the similarity of user test results started to differ significantly from the results produced from the real product. They concluded that if prototypes compromise on the

physical attributes of a design, such as removing the tactile feedback of the buttons, then the performance data was affected. They state that "it is not the level of fidelity that is important but rather the considerations of tangibility and physicality".

Lim et al. [14] conducted trials on a mobile phone in order to understand the effect of fidelity levels on usability data. Three prototypes were tested: the final device, a software representation and a paper prototype. All models picked up major usability issues, but only the final device and software models facilitated the collection of comments regarding the concept's comparison with other products and performance.

In our study, we interpret user data from a trial of a conceptual device as there is no completed device to compare it to. The considerations that have driven the fidelity level and its effects on the physicality of the model have been purely time based. The designer had to decide on the best way to prototype the technical aspects within the allocated time.

User trials were chosen as a means of exploring the effects of fidelity and the resultant physicality on the prototypes. The research of Gill et al. [13] and Lim et al. [14] demonstrate that user trials are an effective way of highlighting design issues by comparing low fidelity models with the final design. Those results gave us the confidence to use similar trials on a conceptual device where there is no 'end product' to compare it to. The aim of comparing the prototypes in this manner was to gather data that enabled a review of the differences in the *way* the prototypes function as each of the prototypes has the same *level* of functionality.
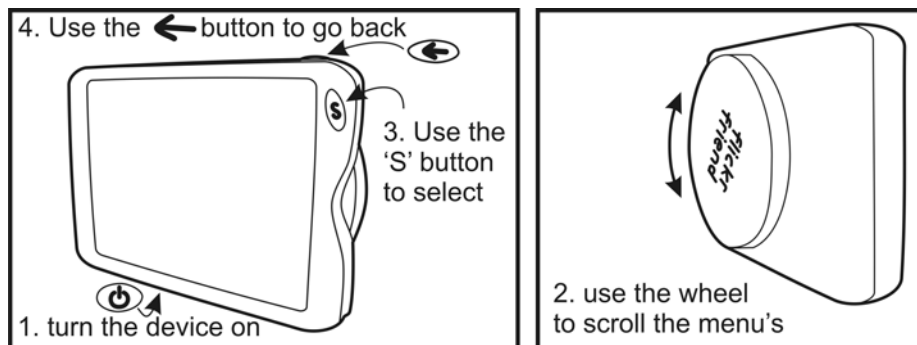
Fig. 1. Different ways of interacting with the device.

3 Our Approach

The trials were conducted on a conceptual device. None of the users had been exposed to the device previously. The concept originated from an undergraduate design brief and was based on the design of a hard drive equipped device offering users the ability to wirelessly view their Flickr [15] web pages and store photos. Flickr is an online photo management and sharing application.

Some initial design work was undertaken in order to develop the physical and digital components of the concept, in order to reach a stage where, in a real design process, an interactive prototype would be the next natural step (see Figure 1 which shows the different ways of interacting with the concept). Each of the resulting prototypes used this initial design work as the starting point, therefore only the time to construct the prototype differed.

3.1 The Resulting Prototypes

'Lowest Level': Time allowed = 4 hours (actual time taken = 3 hours 30 minutes)
Method used: Paper prototyping

Fig. 2. Lowest level prototype: a) the foam model, b) a paper screen c) the trial set up.

As noted earlier, Paper Prototyping is a very simple technique which provides a very fast method for creating low fidelity prototypes. A foam model was constructed to create the physical form to scale. The foam was sanded to produce a smooth finish with white cardboard depicting the buttons and screen (Fig. 2a). For the digital aspect a series of paper screens were created with a small red box to indicate which menu item is active (Fig. 2b). The participant held the physical model, the facilitator changed the screens and adjusted the 'select box' during user trials (Fig. 2c).

'Mid Level': Time allowed = 14 hours (actual time taken = 12 hours)
**Method used: IE System**
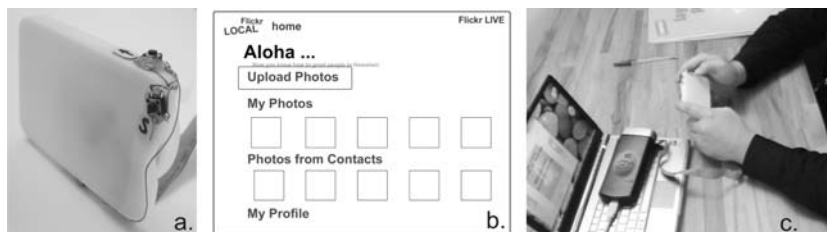
Fig. 3. Mid level prototype: a) the FDM model b) the basic flash interface c) the trial set up.

The IE System was chosen to create the mid level prototype due to the simplicity it offered. The system allows a PC to receive keyboard inputs so that when a user

interacts with a switch in the physical model, the PC will respond to the perceived keyboard input and a keyboard-triggered GUI is activated on the PC. A model was created in a Computer Aided Design (CAD) system and was constructed to scale using a Fused Deposition Modeling (FDM) machine (Fig. 3a). FDM is a rapid prototyping technique where the machine builds the material up layer by layer. A basic menu structure was created in Adobe Flash (Fig. 3b). The Flash animation used keyboard presses activated by off-the-shelf buttons for the screen changes, these were crudely tacked onto the outside of the model and a mechanical rotary dial was glued inside the model for the 'wheel' interaction. For the trial, the physical model was connected, with a cable, to a PC via the IE Unit (Fig. 3c) and the visual feedback was on a desktop monitor.

'Highest Level': Time allowed = 5 days (actual time taken = 5 days)
Method used: IE system and Phidgets
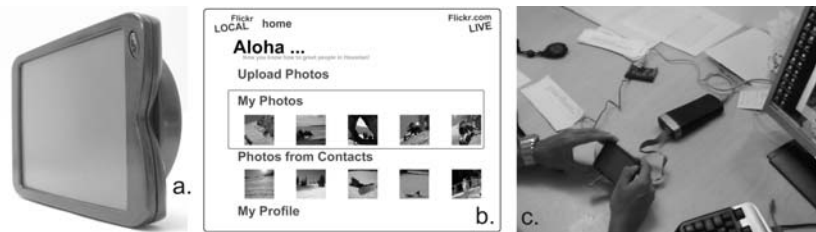
Fig. 4. Highest level prototype: a) the sprayed FDM model b) the flash interface c) the trial set up

The extra time allowed for the highest level prototype was used to develop the following three areas: the prototype was given a realistic finish, the wheel interaction was made to feel smooth and the Flash animation was developed to operate more like the intended design. Again a CAD model was created with design details such as shaped buttons and ports included. Once the FDM model had been made it was sanded and sprayed (Fig. 4a). Dome switches that produce positive tactile feedback with a low profile were used for the buttons triggering the Flash animation through the IE Unit. The smooth feeling analogue dial was an off-the-shelf Phidget component. This reflected the intended physical-digital interaction of the design intent better than the rotary dial used in the mid level prototype. The Flash animation had more realistic menus and a smoother transition between screens (Fig. 4b). For this trial, the physical model needed to be connected through both an IE Unit and a Phidget Interface Kit with wires (Fig. 4c), and the visual feedback was on a desktop monitor.

3.2 Initial analysis of the prototypes created

The resulting prototypes differed considerably and their properties are reviewed in relation to McCurdy et al. [9] five dimensions of fidelity, as shown in Table 1. A

similar technique is applied in Table 2 to analyse the subsequent effects on physicality, which are considered to fall under two areas: the physicality of the device itself (e.g. form, finish, weight) and the physicality of the interaction (feel of the buttons and wheel in this case).

Table 1. Properties of each prototype in relation to the five dimensions of fidelity (McCurdy et al. [9]).

Dimension of fidelity	Driving factors	Lowest level 3 hrs 30min	Mid level 12 hours	Highest level 5 days
Aesthetics	Model material Model finish	**Blue foam** (both material and finish differ considerably from intended design)	**Unfinished FDM** (similar material but finish differs considerably from intended design)	**Sanded & sprayed FDM** (similar material and finish to intended design)
Richness of interactivity	Wheel mechanism	**Free rotating** (similar to intended design but no real-time feedback given)	**'Clunky', clicking mechanism with end points** (very different from intended design but gives real-time feedback)	**Smooth mechanism with end points** (very similar to intended design and gives real-time feedback)
	Buttons	**Cardboard representations** (very different in feel and aesthetics from intended design)	**Switches tacked onto model** (very different to intended design but gives real-time feedback)	**Integrated switches** (very similar to intended design in look and feel gives real-time feedback)
	Screen operation	**Paper screens** (no real-time feedback so very different from intended design)	**Basic Flash animation** (real-time feedback but sketchy interface, differs slightly from intended design)	**More advanced Flash animation** (real-time feedback and graphics similar to intended design)
Depth of functionality	Screen operation	All have identical features enabled, feature will appear 'unavailable' if it is not part of a task		
Breadth of Functionality	Screen operation	All have identical menu structures, the tasks chosen highlighted the breadth of functionality in the intended design		
Richness of Data	Data used	**Sketch data used** (different from intended design)	**Sketch data used** (different from intended design)	**Photos used** (very similar to intended design)

Table 2. Properties of each prototype in relation to the areas of physicality

Area of Physicality	Driving factors	Lowest level 3hrs 30min	Mid level 12 hours	Highest level 5 days
Physicality of the device	Scale	**1:1, made from blue foam with a cardboard screen** (form is very similar to intended design, finish and weight is considerably different)	**1:1, unfinished FDM with screen placement suggested on model (no colour difference)** (form is very similar to intended design, weight and finish are considerably different)	**1:1, finished and sprayed FDM with a colour difference depicting the screen** (form and surface finish is very similar to intended design, weight is different)
	Model material			
	Screen material			
	Weight			
Physicality of the interaction	Wheel mechanism	**Wheel freely rotates** (as intended in design) **with no real-time physical or digital feedback** (extremely different from intended design)	**Mechanism feels clunky and cannot rotate continuously** (considerably different from intended design) **gives real-time physical** (not part of intended design) **and digital feedback** (part of intended design)	**Mechanism feels smooth** (very similar to intended design), **cannot rotate continuously** (not part of intended design) **gives real-time physical and digital feedback** (similar to intended design)
	Buttons	**Buttons are depicted with cardboard and give no physical or digital feedback** (very different to intended design)	**Buttons are off-the-shelf and tacked onto the model** (very different to intended design) **but give real-time physical and digital feedback** (similar to intended design).	**Buttons are integrated dome switches with real-time digital and physical feedback** (very similar to intended design)

4 Method

The set of trials and rating scale used to classify the severity of problems, was based on recommendations by Redish et al. [16]. Participants were divided into three independent groups, with each group using one level of prototype (low, mid or high). Each participant was given a series of 5 scripted tasks [17]:

    Task 1: turn the device on
    Task 2: find a photo on the *Flickr* website
    Task 3: find a friend photo on the *Flickr* website
    Task 4: find a photo from the hard-drive
    Task 5: transfer a photo from a camera

4.1 Structure of the Trials

The following structure was applied to every participant for each of the three prototypes trialed:

i. Participant fills in a demographic questionnaire covering age and gender plus existing technology usage. Note the prototype is not in sight at this stage.
ii. Participant is given a written description of the product.
iii. Facilitator uncovers the model and records if the participant picks it up and her reaction in relation to the fidelity of the aesthetics.
iv. Participant is given the 5 tasks (as described above) to carry out. Facilitator records the time taken for each task and whether the user experienced a success, minor problem, serious problem, or a catastrophe (see Table 3).
v. Participant fills in a questionnaire and is asked to rate certain aspects of their experience with the device.

4.2 The Empirical Study

A pilot study was first carried out with 9 undergraduate participants from the University of Wales Institute, Cardiff (UWIC), which uncovered some problems, including hardware stability issues, and these were then fixed.

The main study was conducted using 48 participants recruited from UWIC staff who have used digital cameras (including cameras on their mobile phones). The participants were divided into three groups of 16, one for each fidelity level, to eliminate possible learning effects. 23 females and 25 males were trialed with ages ranging from 19 to 50, thus an average age of 29. All trials were videotaped for further qualitative analysis.

5 Quantitative Analysis

The quantitative data of interest is the 'performance' data, which shows whether the task was a success, had minor or major problems or was a catastrophe. The data was recorded at the time of each trial based on the criteria shown in Table 3.

Table 3. Description of performance rating and examples

Performance rating	Definition	Examples
Success	Task completed without error	User finds all the correct buttons and menus when needed
Minor problem	Task completed with small error	User goes into the wrong menu, user cannot find a button
Major problem	Task completed with major error/s	User repeatedly tries the wrong menus or buttons
Catastrophe	Task is not completed	User has not completed a task (even if he/she thinks they have), user gives up.

The quantitative analysis was conducted in order compare the results of the prototypes for each of the separate tasks (repeated measures). The performance data was converted into interval data (3 = success; 2 = minor problem; 1 = major problem; 0 = catastrophe) and analysis was conducted using a 3 (prototype level) by 5 (tasks) mixed analysis of variance (ANOVA) with the alpha level set to 0.05.

Figure 5 shows the performance data, a line has been included between the marks to aid interpretation of the graph. No significant overall differences were found between the prototypes. The plots suggest that the prototypes performed similarly for Tasks 1-3, but Tasks 4 & 5 appear to show some differences. Upon further analysis (simple effects) these differences were found to be not significant.
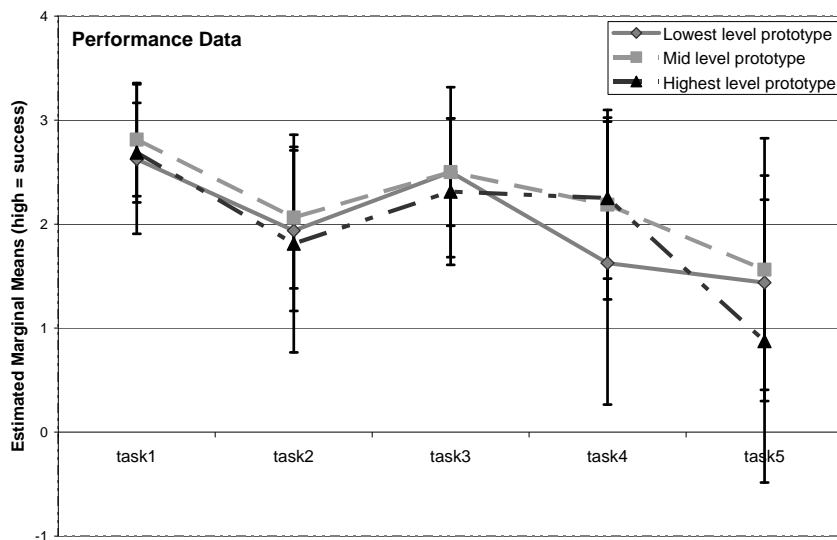

Fig. 5. Performance ratings for each of the 5 tasks as a function of device type

The quantitative data on its own did not reveal any differences, which suggests that neither differences in physicality nor in fidelity have an effect, or that this is not a reliable way of analyzing this effect.

6 Qualitative Analysis

The qualitative analysis was conducted by reviewing the video recordings of each participant after the trials. The qualitative analysis was twofold: firstly, identifying problems that participants may have encountered while performing each task (Part 1 Analysis) and secondly, assessing whether participants were influenced by the fidelity and physicality of the prototypes (Part 2 Analysis).

Part 1 Analysis: This was conducted to find out where participants were having problems performing each task (*types of usability problems*). During the trials, the main errors were observed and noted in a table. Later, while reviewing the video, each

error made by the participant was recorded. If an error had not been listed before, it was added to the table. However, if a participant kept repeating the same error, it was recorded several times, this highlighted particular areas of concern. The errors were then condensed into four problem areas, which we identified as being of hindrance to a user in completing a task. The problems areas are:

    a. Unclear meanings of symbols
    b. Difficulty locating appropriate interface elements
    c. Unexpected feedback from software (mental model mismatch)
    d. Unintentional interaction with software (wanted to interact in a way that was not intended)

Table 4. Number of times usability problems occurred at different prototype level.

Task	Usability problems	Prototype level		
		low	med	high
1	Locating appropriate interface element	9	4	9
	Got it right first time	12	14	11
2	Unclear meanings of symbols	2	1	5
	Locating appropriate interface element	32	13	18
	Unexpected feedback from software (mental model mismatch)	2	6	18
	Unintentional interaction with software	4	0	9
	Got it right first time	2	5	6
3	Unclear meanings of symbols	2	4	1
	Locating appropriate interface element	2	2	8
	Unexpected feedback from software (mental model mismatch)	12	2	16
	Got it right first time	10	9	8
4	Locating appropriate interface element	2	0	2
	Unexpected feedback from software (mental model mismatch)	22	17	12
	Unintentional interaction with software	6	4	4
	Got it right first time	7	8	6
5	Locating appropriate interface element	0	8	3
	Unexpected feedback from software (mental model mismatch)	23	18	16
	Unintentional interaction with software	1	4	8
	Got it right first time	3	4	4

Table 4 shows the number of times users encountered usability problems for each task at different prototype levels. The results that are of particular interest are those that differ across the prototypes. So, for example, during Task 2 there were 2 problems recorded by the lowest level prototype due to unexpected feedback from the software but. The same task resulted in 18 problems for the highest level prototype. Other notable results are again for Task 2 where users could not locate the appropriate interface 32 times for the lowest level, 13 times for the mid level and 18 times for the highest level. Interestingly for Task 5, there were 0 problems for the lowest level prototype in locating the appropriate interface elements, but 8 problems for the mid level and 3 problems for the highest level.

Further analysis of the problems related to Task 2 suggests that users of the lowest level prototype had so much trouble identifying the correct interaction (32) that there

were very few mental model mismatch issues (2). Compare this to the highest level prototype, where users were able to find the interaction better (18 errors), but they had difficulty with the mental model of the device (18). The inability to identify the correct interaction could arise either because of a lack of understanding of the symbols (which were the same across the prototypes) or a complete misunderstanding of the results of that form of interaction. The mid level prototype instead has the lowest number of problems related to 'identifying the interaction' (13) and an average range of problems with the mental model (6). So what could be the reason behind these problems? From Table 2, we can see that the lowest level prototype has no tactile feedback on pressing the buttons (just the facilitator moving a screen), while the mid level prototype has very pronounced buttons that give both tactile and on screen feedback, and the highest level prototype has more subtle visual properties with subtle tactile feedback plus on screen feedback. The number of problems linked with locating the appropriate interface element in Task 5 could have arisen due to the same issues as in Task 2, in other words, users of the lowest level prototype had already made so many mistakes early on that they are less likely to make mistakes in the later tasks, unlike users of the highest level prototype who are still experiencing problems even in the later tasks.

Part 2 Analysis: This was undertaken to assess whether participants were affected by the fidelity and physicality of the prototypes based on the related comments made, for example, 'wheel mapping not natural'. A similar recording procedure was followed as in Part 1 Analysis using the errors noted during the trials plus the video review. The comments were then sorted and the ones related to the following areas were selected:

1. physicality of the device (e.g. size in the hand, screen position and size)
2. physicality of the interaction (e.g. the button is in the wrong place, how the wheel feels etc.)
3. feedback about the design and idea in general

The results are shown in Table 5. The general feedback on the design and concept is roughly the same across the prototypes. The lowest level prototype seems to differ in the number of comments about both the physicality of the *device* (22 at the lowest level compared to 13 at the mid level and 16 at the highest level) plus the physicality of the *interaction,* 42 at the lowest level compared to 52 at the mid level and 57 at the highest level. These results suggest that the test was set up in a way that entices generally attracted more comments about the physicality of the interaction rather than the physicality of the device. However, the lowest level prototype received more comments about the physicality of the device unlike the mid and highest levels. This could be because the physicality of the interaction was so far removed in the lowest level prototype from that intended, hence it was harder for users to judge this aspect of the design and as a result, they made more comments about the physicality of the device itself.

Table 5. number of comments related to the physicality and fidelity at different prototype levels.

	Lowest level	Mid level	Highest level
Physicality of the device	22	13	16
Physicality of the interaction	42	52	57
General feedback on the design and	19	17	18

7 Discussion

Each of the prototypes created represented the same design intent and enabled the same functionality. Time constraints governed the fidelity level and each prototype was tested for usability and physicality issues. The prototypes needed to convey enough information to the users so they were able to get a feel for the design intent of the product. The initial hypothesis was that fidelity and subsequently physicality would have an effect on the users understanding of the product and therefore user feedback and usability would be affected. The analysis of the user trials showed the following results:

1. All prototypes achieved similar results for the performance test.

There was in fact little difference in performance across the prototypes with different fidelity levels (which would seem to agree with the research by Lim et al. [14]). This in itself is an important result showing that in the early stages of the design process, the fidelity level might not have a significant impact. Despite the mid level prototype being physically different from the intended design in a number of seemingly important ways (the wheel clicked, could not rotate 360° and felt very 'clunky'), it still produced valid feedback about the concept. Furthermore, the mid level prototype took less than half the time to build compared to the highest level prototype. Even the lowest level paper prototype seemed to produce usability data in line with the higher fidelity ones.

2. Users of the mid and highest level prototypes, with real time tactile and digital (on screen) feedback, had fewer problems in locating the appropriate interface element.

Even when all the prototypes had the same graphical symbols, the lowest level prototype users had a lot of problems identifying the appropriate interface elements. This may be because many users worked out what interactions did by 'experimenting' with them instead of understanding the symbols and thus made their decisions based on the feedback they received. This approach was only supported by those prototypes that had real time feedback, whereas the lowest level paper prototype required the facilitator to find the correct feedback and change the screen.

3. Users of the mid and highest level prototypes had more problems with the mental model of the device early on in the trial whereas the lowest level prototype users encountered these issues later on in the trial.

This is an unexpected outcome. Table 4 shows that, even after completing 4 tasks, users still made errors due to a mental model mismatch for task 5. Users who had real time tactile and digital feedback from their interactions had more difficulty in understanding how the device worked. The most likely explanation for this is that users of the lowest level prototype were so distracted by not locating the appropriate interface element that this overshadowed their understanding of the device. By the end of the trial, users of the lowest level prototype were having less problems locating the interface element but more difficulty in understanding the device (their mental model). This could possibly be due to the users' inability to fully engage with the device and therefore following a 'more luck than judgment' approach.

4. The mid and highest level prototypes gave more feedback about the physicality of the interaction.

This was not unexpected as in order to get valid feedback about an interaction, one needs to approximate the intended interaction, which the lowest level paper prototype did not facilitate.

8 Conclusion and Further Work

This paper has reported on a preliminary investigation into the effects of physicality and fidelity on the prototypes used for front end product design development. The trials suggest that there is no effect of fidelity at the early stage of the design process in terms of user performance, however a deeper analysis is required. As expected, the qualitative analysis showed that prototypes that gave real time interaction and feedback allowed users to get a more realistic appreciation of how the device worked, and also generated more useful comments about how the device feels to hold and to interact with.

From these results, we can draw that for the initial exploration of a design idea, very low fidelity prototyping is a fast and low cost method of getting reliable feedback. On the other hand, if more specific feedback about the intended design and interaction is required, then a prototype that can produce immediate feedback is essential. However, there are many more factors at play and these need to be researched further to inform design guidelines in relation to the needs of the early design process.

The nature of physicality seems to have an impact on the user trials of these prototypes, but a very in-depth analysis had to be carried out to tease out these effects. It would be more useful if such effects could be found and explored using faster quantitative analysis. Further work needs to be undertaken to explore how these effects of physicality can be tested in a quantifiable way and therefore fully explore the wider implications for designers in practice building.

9 References

1. Schrage, M.: Serious Play: How the World's Best Companies Simulate to Innovate. Harvard Business School Press (1999)
2. Snyder, C.: Paper Prototyping: The Fast and Easy Way to Design and Refine User Interfaces. Morgan Kaufmann (2003)
3. Hartmann, B., Klemmer, S.R., Bernstein, M., Abdulla, L., Burr, B., Robinson-Mosher, A., Gee, J.: Reflective physical prototyping through integrated design, test, and analysis.

Proceedings of the 19th annual ACM symposium on User interface software and technology. ACM, Montreux, Switzerland (2006)

4. Greenberg, S., Fitchett, C.: Phidgets: easy development of physical interfaces through physical widgets. Proceedings of the 14th annual ACM symposium on User interface software and technology. ACM, Orlando, Florida (2001)

5. Burleson, W., Jensen, C.N., Raaschou, T., Frohold, S.: Sprock-it: a physically interactive play system. Proceedings of the 6th international conference on Interaction design and children. ACM, Aalborg, Denmark (2007)

6. Gill, S.: Developing Information Appliance Design Tools for Designers. 1st Appliance Design Conference, Bristol, UK (2003)

7. Sefelin, R., Tscheligi, M., Giller, V.: Paper prototyping - what is it good for?: a comparison of paper- and computer-based low-fidelity prototyping. CHI '03 extended abstracts on Human factors in computing systems. ACM, Ft. Lauderdale, Florida, USA (2003)

8. Virzi, R.A., Sokolov, J.L., Karis, D.: Usability problem identification using both low- and high-fidelity prototypes. Proceedings of the SIGCHI conference on Human factors in computing systems: common ground. ACM, Vancouver, British Columbia, Canada (1996)

9. McCurdy, M., Connors, C., Pyrzak, G., Kanefsky, B., Vera, A.: Breaking the fidelity barrier: an examination of our current characterization of prototypes and an example of a mixed-fidelity success. Proceedings of the SIGCHI conference on Human Factors in computing systems. ACM, Montreal, Quebec, Canada (2006)

10. Ghazali, M., Ramduny-Ellis, D., Hornecker E., Dix A. (eds.): Proceedings of the First International Workshop on Physicality, Physicality 2006, (6-7 Feb. 2006) Lancaster University, UK (2006)

11. Ramduny-Ellis, D., Dix, A., Hare, J., Gill, S.: In: Ramduny-Ellis, D., Dix, A., Hare, J., Gill, S. (eds.): Proceedings of the Second International Workshop on Physicality, Physicality 2007, (2-3 Sep. 2007). UWIC Press. Lancaster University, UK (2007)

12. Norman, D.A.: The next UI breakthrough, part 2: physicality. Vol. 14. ACM (2007) 46-47

13. Gill, S., Walker, D., Loudon, G., Dix, A., Woolley, A., Ramduny-Ellis, D., Hare, J.: Rapid Development of Tangible Interactive Appliance: Achieving the Fidelity/Time Balance. International Journal of Arts and Technology (IJART) Special Issue on Tangible and Embedded Interaction **1** (2008) 309-331

14. Lim, Y.-k., Pangam, A., Periyasami, S., Aneja, S.: Comparative analysis of high- and low-fidelity prototypes for more valid usability evaluations of mobile devices. Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles. ACM, Oslo, Norway (2006)

15. Flickr: www.flickr.com. (2004)

16. Redish, J., Bias, R.G., Bailey, R., Molich, R., Dumas, J., Spool, J.M.: Usability in practice: formative usability evaluations - evolution and revolution. CHI '02 extended abstracts on Human factors in computing systems. ACM, Minneapolis, Minnesota, USA (2002)

17. Rubin, J.: Handbook of Usability Testing. John Wiley & Sons (1994)