Comparing user interaction with low and high fidelity prototypes of tabletop surfaces

6 authors, including:

Some of the authors of this publication are also working on these related projects:

ALADIN View project

Algorithmic Mediation in Academic Social Systems (AMASS) View project

Comparing User Interaction with Low and High Fidelity Prototypes of Tabletop Surfaces

**Jan Derboven[†], Dries De Roeck[†], Mathijs Verstraete[†], David Geerts[†],
Jan Schneider-Barnes[‡], Kris Luyten[‡]**

[†]Centre for User Experience Research (CUO)
IBBT/K.U.Leuven
Parkstraat 45, 3000 Leuven, Belgium
{firstname.lastname}@soc.kuleuven.be

[‡] Hasselt University – tUL – IBBT
Expertise Centre for Digital Media
Wetenschapspark 2, 3590 Diepenbeek, Belgium
{firstname.lastname}@uhasselt.be

ABSTRACT

This paper describes a comparative study between the usage of low-fidelity and a high-fidelity prototyping for the creation of multi-user multi-touch interfaces. The multi-touch interface presented in this paper allows users to collaboratively search for existing multimedia content, create new compositions with this content, and finally integrate it in a layout for presenting it. The study we conducted consists of a series of parallel user tests using both low-fidelity and high-fidelity prototypes to inform the design of the multi-touch interface. Based on a comparison of the two test sessions, we found that one should be cautious in generalising high-level user interactions from a low towards a high-fidelity prototype. However, the low-fidelity prototype approach presented proved to be very valuable to generate design ideas concerning both high and low-level user interactions on a multi-touch tabletop.

Author Keywords

Tabletop, multi-touch, prototyping, fidelity, content creation, design.

ACM Classification Keywords

H5.2 User Interfaces: Input devices and strategies

INTRODUCTION

In the field of HCI, methods for prototyping computer interfaces have quite a history [3, 7, 10, 20, 23]. For testing most graphical user interfaces, simulating interactivity can be done using low-fidelity prototyping methods: these are simple visualisation techniques of design ideas at very early stages of the design process, whose development does not need much time, such as paper prototyping. Low-fidelity prototyping methods typically involve "low-fidelity materials" that are different from the final product, such as paper, cardboard, etc. [20, 21]. High-fidelity, functional interactive prototypes are built at a later stage of the user-centred design process, and typically do use materials that you would expect to be in the final product, producing a prototype that looks and behaves much more like the final product [21]. In multi-touch technology and interaction, however, directness and visual interactivity is at the core of the experience and the interaction [16]. Because of that, it becomes hard to apply Wizard of Oz-like techniques to multi-touch directly, as problems with delay or obscured parts of the interface make it difficult for the 'wizard' to quickly and accurately respond to user interactions. Some existing multi-touch developments deal with this by means of a 'design by doing' approach, where ideas are implemented straight away and tested afterwards, without having a profound notion of the usefulness or chance to succeed at the start of the development [8]. From a developer's point of view, this is an understandable approach since the technical complexity of multi-touch hardware is relatively low. The downside of this is that implementing systems straight away is error-prone, and thus costs more time and money than creating low-fidelity prototypes. The required effort to update a system after it has been developed is significantly higher compared to creating several, rapid iterations of low-fidelity prototypes [23]. Furthermore, when implementation is started immediately, HCI practitioners coming from a non-technical background are experiencing difficulties to fully participate in the process of creating prototypes [20], and exploring the potential of the new interaction paradigms.

To overcome these limitations, and stimulate the use of low-fidelity prototypes in the design of multi-touch tabletop interaction, we compare a low-fidelity and a high-fidelity prototype for the design of a multi-touch user interface.

The paper is structured as follows. First, we will survey research in the area of prototyping, mostly focusing on multi-touch surfaces. We will then present the MuTable platform, which is used as a case for comparing low-fidelity and high-fidelity prototyping. The two test set-ups will be described in detail, before presenting the results which focus on the differences and similarities on different aspects

of the prototypes. We will conclude with findings about the value of multi-touch prototypes in various phases of development.

RELATED WORK

Low-fidelity prototypes can be used for different purposes. In a HCI context, paper prototypes are mostly used as tools for the evaluation of design/usability failures or successes [6]. The method is used to validate a prototype of which the findings are applied in a final product or application [1, 15]. Another approach to paper prototyping is proposed by Lim, Stolterman and Tenenberg [7], i.e. the use of a low-fidelity prototype as an explorative way to look at design ideas and as a way to think about design alternatives. Scott, Carpendale and Inkpen [19] used paper prototyping to categorize the type of interactions people have when they work collaboratively at a physical table. Findings of their study concern high-level interactions starting from a paper prototype, and do not focus on the relation with an actual multi-touch application. Research by Al Mahmud et al. [1] shows that interactivity and visualization is the core of a multi-touch experience, which is why the physical design of a low-fidelity prototype for multi-touch is something to be approached with caution. Therefore perhaps, few efforts have been made to experiment elaborately with 'low-fidelity', physical interaction in a multi-touch development context. Examples are [14, 19, 24]. Both Piper [14] and Terrenghi [24] use a low-fidelity multi-touch oriented prototype in a focussed context, where the functionality of the tested application is relatively narrow. Furthermore, it is not always clear whether low-fidelity prototype studies allow for generalization to high-fidelity multi-touch products. Terrenghi [24] presents a study investigating this. In her study, a comparative analysis was made between a high and low-fidelity version of single-user puzzles and photo sorting tasks, focusing mainly on interface metaphors and lower-level interaction (such as one-handed vs. bimanual interaction).

The approach taken in this paper focuses on higher-level interactions, such as interpersonal interaction between users. Specifically, the similarities and dissimilarities between low-fidelity prototypes and their high-fidelity counterparts in collaborative content-creation tasks are investigated.

THE MUTABLE PLATFORM

Our tabletop surface setup, conveniently called MuTable in the remainder of this paper, was used in a variety of public spaces, including museums, public events, libraries, and schools. While a lot of multi-touch applications typically focus on viewing video, browsing photographs and playing music, MuTable focuses on productivity tools for collaborative content creation. The usage scenario central in this paper is one where adolescents are asked to collaboratively use a tabletop interface for school-related work. Students are able to search for content (text documents, pictures, movies…), and integrate this content

in a presentation. The scenario integrates searching and browsing for content with the creation of new content.

A lot of work has already been done in the field of collaborative multi-touch applications, both on the side of detailed interaction design (for instance, gesture research as described in [12, 17, 25, 26]), as on a higher-level plane of application design and user collaboration [18]. Applications range from (collaborative) picture browsing and sharing [2, 5, 13], to gaming [1, 8, 15], storytelling [9, 22] and even specialized applications for hotels, restaurants and healthcare (e.g. the Sheraton Hotel, Resto'Touch, and Patient Consultation showcased in [11]). However, significantly less attention has been paid to productivity tools and specifically content *creation* on multi-touch surfaces. The specific focus of the user tests described in this paper is on researching how people behave and interact with each other and the multi-touch work surface at a high level, while they are working on a *creative* assignment.

The MuTable was conceived to be a multi-touch surface that is positioned horizontally, or at a small angle, convenient for people to work at while standing up. The surface was specifically intended for co-located collaboration when standing at one particular side of it, so all users face the screen in the same direction, unlike some other multi-touch hard- and software designed for access from all four sides of the surface (e.g. Microsoft Surface). This consideration also played its part in UI design, as the software was designed with users in mind standing in one specific orientation at the table. This setup makes the UI design of the MuTable software very suitable for touch tables that are placed against a wall, or that are wall-mounted. When the table is positioned at a small angle in one direction, the table position will also be ergonomically better to work at, compared to a completely horizontal surface. Moreover, the MuTable design incorporates a full-screen presentation mode in which slides created on the table can be presented. This functionality is aimed at wall projection, or wall-mounted tables that can be tilted from a small angle (in 'working mode') to a vertical position, to serve as a presentation surface ('presentation mode').

METHOD

In our study, we performed two series of user tests: one focusing on low-fidelity prototyping and another on high-fidelity prototyping. The two observation series, which are described in more detail below, were very similar in set-up. Different groups, equal in size and of the same age, worked on the same assignment. The main difference between the two tests was the material the participants worked with: the 'low-fidelity' groups created a presentation with physical everyday objects (paper, scissors, etc.), while the 'high-fidelity' groups created a similar presentation using only digital content available on a multi-touch table.

When testing low-fidelity prototypes, and comparing them to their high-fidelity counterparts, it is important to choose the right prototyping focus. Arnowitz et al. [3] identify

several 'prototype fidelity levels': a prototype does not have 'just one characteristic of fidelity', but it has several fidelity levels. Arnowitz et al. distinguish between levels such as information design, interaction design, navigation model, visual design, etc. However, the focus aimed for in the MuTable low-fidelity prototype doesn't fit neatly into one of the levels described by Arnowitz et al. The aim is to look at a general level of interpersonal interaction and collaboration, and general interaction with the work surface and the material on it. In the test setup described below, the similarity in participants, group size, assignment, and tools should allow for a high-level comparison between the two series of user tests. The test setup does not, however, allow for detailed comparisons of specific interaction techniques, as the low-level interactions of handling scissors and picking up or folding paper in the low-fidelity prototype are not comparable to handling digital information in the high-fidelity prototype. As also found in related work, such low-level comparisons would require different prototype setups, in which it can be more difficult to draw conclusions about high-level interaction.

Low-Fidelity Tests

Prototype Design
The low-fidelity prototype of the interactive table involved a table at which test users were asked to collaborate using only real-world, physical materials. The long side of the table (120x80x92 cm - LxWxH) was placed against the wall, in order to match the requirements of the envisioned touch table. At the start of the observations, only 3 sheets of A3 paper were on the table: these sheets were to be used to create a presentation on a given subject.

In order to allow for a comparison between the low-fidelity and the high-fidelity prototypes, the former prototype's fidelity was increased by imposing a small number of restrictions. The materials to be used for the creation of the presentation were available in a separate material repository, which was hidden behind a curtain. The participants received a list of materials available for use in the presentation – they had to ask the test facilitators to bring them the material they needed. The repository included documentation materials, a typewriter to type text, a 'recycle bin', and materials to paste (glue, tape, etc.), to cut (cutter knife, scissors, etc.), and to draw. In order to simulate screen estate usage, every material group (e.g. a set of scissors and knives) was in a separate container that had to be placed on the table. Participants were not allowed to stack the containers, or put them on the floor: all materials had to be put on the table, or given back to the observers.

Since the materials were not immediately available, participants could not really anticipate what to expect when asking for specific work material (Figure 1). Test participants could ask the test observers at any time to bring them work materials from the repository, or to return these materials to the repository. In this way, the prototype

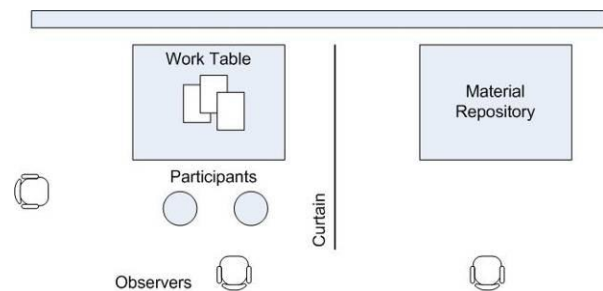mimics a user gradually exploring a system during first time use, opening and closing functionality.

Figure 1: Low-fidelity test setup

Test Design
For the test with the low-fidelity MuTable prototype, 16 adolescents aged 16-17 were recruited (7 female, 9 male participants). They participated in small groups (2 groups of 3, 5 groups of 2), in which all members of the groups already knew each other before the test: they were friends, classmates, or family.

The participants were asked to work at the prototype of the interactive table, standing up (see figure 2). Their task was to create a presentation in 45 minutes on the life and works of Leonardo da Vinci. This topic allowed the participants to create a presentation about the aspects of Leonardo they thought most interesting (arts or science), and it resembles a school assignment, as stipulated in the MuTable use scenario.

Figure 2: Low-fidelity table usage

For analysis purposes, the test sessions were recorded using two video cameras. One of the test observers gave the assignment to the participants, in which the participants were asked to create a presentation for a travelling exhibition on Leonardo da Vinci. When finished, the participants presented their work and talked about the contents of their presentation. Before starting the test, participants were given a pre-test questionnaire asking about their background; after the test, they were given a post-test questionnaire asking how the assignment went. At the end of the test session, the participants received an incentive.

High-Fidelity Tests

Prototype Design
The high-fidelity prototype offers similar functionality for creating presentations as the low-fidelity prototype discussed in the previous section. In general, the high-fidelity MuTable platform offers a number of separate, small applications such as a typewriter, a presentation creation tool, a file browser, and a drawing application (see figure 3). The functionality is available on a surface of 110 cm x 55 cm, with a screen resolution of 1280x800 pixels. This allows two to four persons to collaborate at the MuTable in a comfortable way, depending on the specific tasks to be completed.

All applications can be accessed through a central, circular menu, which can be personal (one menu per user) or common to several users: no restrictions have been imposed regarding 'ownership' of content. The applications can be used together to create and search for content, and (re-)use that content in a presentation. The applications especially facilitate research on a specific subject, and creating new content (text, a presentation) on that subject. The high-fidelity MuTable applications are the digital counterparts of the material used in the low-fidelity setup. Test participants in the high-fidelity tests, like the low-fidelity participants, could also do research and compile a presentation by reading documents, typing texts, using pictures, dropping unwanted material in a recycle bin, etc. In this way, both prototypes were very similar in their high-level functionality, regardless of lower-level, specific differences (such as the use of specific interaction methods, like gestures).
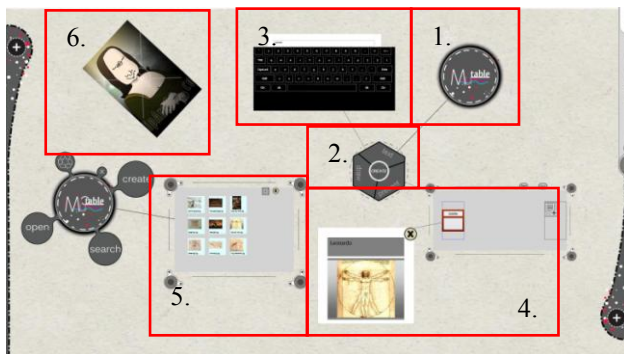

Figure 3: Main MuTable functionality, including 1. a central circular menu, 2. a Create submenu, 3. a typewriter, 4. a presentation creation tool (with a slide opened), 5. a file browser, 6. a piece of content (picture) opened.

Test Design
For the tests with the high-fidelity MuTable prototype, a test design similar to the low-fidelity test design was used. 16 adolescents aged 16-18 were recruited (9 female, 7 male participants – all were different participants from the ones that used the low-fidelity prototype). They participated in small groups (2 groups of 3, 5 groups of 2), in which all

members of the groups already knew each other before the test: they were friends, classmates, or family.

One of the test observers gave the assignment to the participants, which was the same assignment used in the low-fidelity tests: creating a presentation for a travelling exhibition on Leonardo da Vinci in 45 minutes. The test itself consisted of some ten minutes in which the participants were allowed to explore the table themselves. After some ten minutes, most participants arrived at a point where they needed some extra information and guidance on specific gestures and the functionality of the platform. To provide the test participants with extra information after their initial exploration, the complete MuTable functionality was explained to the participants in a short demonstration. Afterwards, they were allowed to continue working on their presentation for approx. 30 minutes. For analysis purposes, the test sessions were recorded using three video cameras. After completing the test assignment, participants presented the work they had created. Before starting the test, participants were given a pre-test questionnaire asking about their background. After the test, they were given a post-test questionnaire asking how the assignment went. At the end of the test session, the participants received an incentive.

Comparing Low-Fidelity and High-Fidelity Tests
The observations and user feedback from both the low-fidelity and the high-fidelity test sessions were processed separately. In a first round, the test sessions were discussed and the common, recurrent themes/observations were logged. Based on this limited set of themes, the test sessions were reviewed a second time, fitting all relevant observations into the predefined themes. Afterwards, the observations were compared to each other, in order to find parallels in user behaviour between the two test settings.

RESULTS
In this section, the results of the test sessions are divided into a number of key areas regarding high-level user interaction. The test results of both test sessions will first be discussed separately. After the separate discussion, an overall evaluation and comparison of the high and low-fidelity test results will be made.

Table Layout

Low-Fidelity Table Layout
Low-fidelity test participants tended to divide the table into several zones, similar to those described by Scott et al. [19] – for convenience sake, we will use the same terminology. One type of zone observed during the low-fidelity tests was the personal zone. This zone, typically the zone closest to a specific participant, was used only by that individual participant, and was primarily used for presentation creation, and to store ready-to-use documentation materials. The second type of zone was the group zone: this zone was used for common storage of the tools and unused documentation materials. Although some test groups clearly

were messier in their work than others, this division provided a structure for the test participants in which they knew where to look for a specific tool or content piece, even if one of the other participants used it last. E.g. the pair of scissors was always returned to the same storage/group zone of the table, regardless of which participant used it.

The personal work zone consisted of the presentation being created, and numerous pages and scraps of content lying around, ready to be used in the presentation. The fact that the work zone consisted of current work materials and other content ready to be used, is consistent with the observations of Scott et al. [19] about storage zones, which states that 'the location of a storage territory appeared to influence who utilized the resources contained within it'. Shared items were indeed kept in the group storage zone, while 'stored materials' ready to be used by one participant were kept close to that participant (see figure 4).

However, while Scott et al. only mention the use of piles to store content [19], the MuTable low-fidelity observations adds content layering to this framework. Instead of creating neat piles of content, participants tended to work on top of one or several layers of ready-to-use materials, randomly scattered on the table. In this way, most test groups effectively had more than one 'layer' in their personal zones. The upper layer consisted of the presentation pages, and some tools they were actively using; the lower layer (often invisible, covered with the upper layer) consisted of content scraps lying around, waiting to be used. The participants lifted the upper layer in order to browse the underlying layer of content scraps.
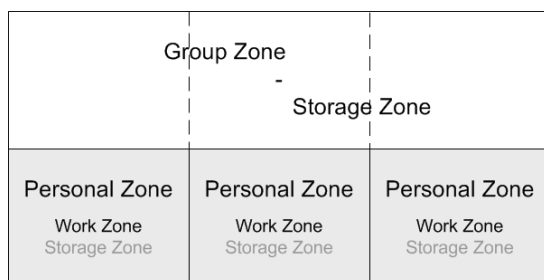

Figure 4: Table layout emerging in most low-fidelity tests (three participants). The high-fidelity layout was often less clear-cut.

High-Fidelity Table Layout
The MuTable high-fidelity prototype didn't impose any fixed table 'layout'. As in the low-fidelity prototype, applications as well as content could be placed anywhere on the table, and they could be resized (and, for content, also rotated) to any size. In the MuTable software, there are no defaults for window or content positioning, such as in graphical user interface (GUI) operating systems (e.g. Minimize and Maximize buttons in MS Windows, Split and Arrange All options in individual applications). Space and screen estate was to be divided by the test participants

themselves, similar to the low-fidelity tests, in which participants also were free to divide the table according to their wishes. All test groups more or less naturally took a position between the two extreme approaches to table layout: a more fixed, clean layout emerged naturally from a task division between the test participants. Without a clear task division, the table layout tended to be messier.

- **Messy table layout.** A messy table layout was in most cases the result of unclear task divisions between the test participants. Test groups that tended towards this side of the continuum mostly got annoyed with this arrangement after a while, commenting on it: '*Shouldn't we close some stuff first?*', and '*Maybe we should give the Recycle Bin a permanent place somewhere*'.

- **Fixed table layout.** This table layout mostly emerged from a more or less clear task division, e.g. one test participant typing text and creating slides, and another test participant searching for content such as images and documents. Four out of seven test groups evolved towards a more or less stable table arrangement in this way. For instance, one test group placed the keyboard in the lower left corner of the table, the presentation creation tool in the lower right corner, and the file browser in the upper right corner. The division left-right often demarcated the individual users' work spaces, while users kept the functionality they were working on most actively closest to them, and other functionality further away (division lower-upper half of the table – see also the section on interaction among test participants).
One test group, however, also used a more or less stable table arrangement: they placed the typewriter centrally in the lower half, and the presentation creation tool centrally in the upper half. In this way, the table layout didn't really support simultaneous work: instead, they collaborated on the same tasks, the two test participants alternately working and observing.

Table Layout: Comparison
The table arrangements made by test participants during the low-fidelity and the high-fidelity tests were often very similar, in terms of global arrangement. Nearly all test groups divided the table in smaller vertical zones right in front of them, which were the users' personal zones. Apart from that, users also divided the space into horizontal zones: one personal work zone with material they were actively working on, and one common storage zone with material not used as actively, but still ready to be used.

Clear table layouts were somewhat more obvious in the low-fidelity tests, while layout in the high-fidelity prototype tended to be somewhat more 'fluent'. It is striking that in the high-fidelity prototype, a clear task division eventually lead to a more fixed layout, and an unclear task division lead to a messier table layout. Most test groups started out exploring with a messier layout – some groups later evolved into a more structured one. In the low-fidelity

prototype, this effect of task division was not that clear: the same table layout emerged, whether the participants had a clear task division or not. This leads to the conclusion that working with tangible materials that can be picked up and 'layered', a suitable table layout emerges more fluently. It seems that the 2D, flat surface of the multi-touch table is somewhat limiting in this respect, an issue also explored by Terrenghi et al. [24]. Test groups that tried to work in an organic way, but weren't able to layer the material in front of them, often ended up with an unstructured mess – the more successful groups were 'forced' into a more structured layout due to the two dimensionality of the content on the table. To make up for this, more explicit guidance on how to effectively divide the table into zones would be a good user support addition.

Interaction among Test Participants

Low-Fidelity Participant Interaction
When observing participant interaction, two general kinds of interaction could be distinguished, similar to the observations described in Gutwin et al. [4]. Test participants switched fluently between working individually on separate tasks and working together (e.g. working together creating content, in which one participant reads a piece of text out loud and the other one types out relevant items on the typewriter). Most of the time, participants worked individually. Even when they were all working on similar tasks, such as placing content on the presentation pages, participants tended to work individually (e.g. by each working on a different slide), merging their results afterwards. This phenomenon is described by Gutwin et al. [4] as 'coupling': working individually in a group before 'requir[ing] discussion, instruction, action, information, or consultation with another person'.

High-Fidelity Participant Interaction
The interaction among the high-fidelity test participants generally went smoothly, apart from some minor incidents, such as users getting in each other's way while working on separate tasks (these incidents will be discussed in the section on territoriality).

In general, the same two kinds of interaction as described in the low-fidelity setting were observed: alternately working individually on separate tasks and working together. However, a distinction was seen between parallel, individual work, and turn-taking.

- Parallel, individual work. A minority of test users worked individually during the entire assignment, with participants working constantly on parallel, separate tasks. This parallel collaboration implies that all test users had a personal zone on the multi-touch table, in which they worked separately. When one participant found some content that could be used by another participant, it was exchanged in a fluent, natural way. Like in the low-fidelity tests, individual work was interrupted by episodes of merging results.

- Turn-taking. Turn-taking collaboration was observed most often, interwoven by episodes of collaboration. Users didn't work individually, but took turns in actually interacting with the table, e.g. one user was just watching the other one interact with the table, or one user was reading text from a content PDF, while the other user was actually typing the text. These shifts between participants occurred mostly guided by negotiating, and comments such as 'Let me try!' or 'Can I type some text now?'

The prevalent turn-taking behaviour obviously has an important effect on the results shown in graph 1. For instance, there was a 3-user test group in which the users didn't really take turns, but were actively working most of the time, dividing the table into 3 work zones (one zone per test user). With this in mind, it is understandable that two users of this group rate the table as 'a bit too small', and only one user as 'big enough, no room to spare'.

The opposite observation is also true. Three groups out of seven mostly took turns in actively using the table; all users in these groups but one reported that the table was 'big enough, with some space left', even though one of these groups had the maximum test group composition of three users.
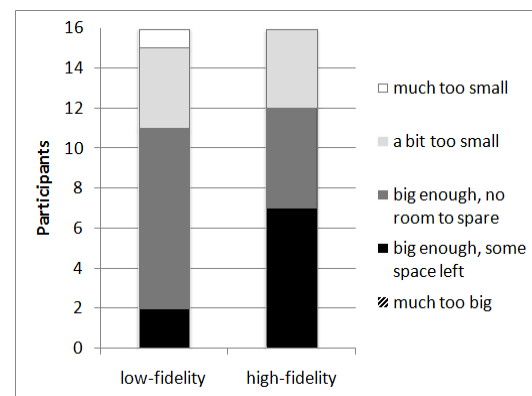
Graph 1: Table size ratings in the post-test questionnaire.

Participant Interaction: Comparison
Although overall, the participants in both sessions showed similar behaviour of alternating individual work with working together, an important distinction in the individual work was observed. While in the low-fidelity setting, people most of the time worked individually, in parallel, the high-fidelity setting showed people taking turns in working. In most groups, there was an active role ('worker') and a passive role ('onlooker'), with participants constantly switching between the two roles. Since participants constantly had to work on the same surface, they tended to let one participant work at a task alone before doing work themselves. Only after some moments of on-looking, other participants started working themselves. 'Tangible' work with 3D objects, as in the low-fidelity prototype, allowed more for individual work, without anyone looking. 2D interaction, on the other hand, often created situations in

which participants tended to let another participant work, without 'interfering' by starting to interact themselves. In this view, reducing interference would only be possible by increasing table size quite drastically.

This behaviour impacted the way the table size was judged. With participants taking turns in working, the high-fidelity table was rated as 'big enough, with some space left' by most participants. In contrast, most low-fidelity participants – working in parallel most of the time – rated their table as 'big enough, no room to spare' (see graph 1).

Territoriality

Low-Fidelity Territoriality
In the low-fidelity test, users collaborated quite naturally. Participants exchanged material, passed material on to other participants, and often reached across the table to find the right tools to work with. Participants didn't really get into each other's way while doing this. These findings are similar to the results described in [19], where participants partitioned their interactions 'with little to no verbal negotiation'.

Furthermore, users frequently exchanged tools and content material among each other. Participants passed on tools and showed photos and documents to each other. Larger tools, such as a typewriter, were not exchanged (which we thought was most likely due to physical properties such as weight). Instead, users changed places behind the table to type text on the typewriter. In general, users collaborated quite fluently, reached over to pick up the material they needed, and switched positions, e.g. to use the typewriter.

High-Fidelity Territoriality
As already hinted at in the previous section, test participants often took turns working. Consequently, when only one user is working at the table, the user encounters no issues regarding the demarcation of his or her work territory. However, most groups worked simultaneously on more or less separate things at some point in the presentation creation (e.g. typing and browsing). This also means that the individual participants in these groups had their 'own' functionality or content they were working on at any given moment. However, since there was no strict table division imposed, individual users' work sometimes got in the way of other participants' work. This was especially the case when resizing content and functionality on the table – not so much when moving functionality or content around the screen. In the last case, most participants appeared to 'respect' each other's territory. When resizing content or functionality to get in the way of other users, remarks about it were made immediately: 'Hey, this is MY part of the table!', 'Keep your things in your own part, man!', 'You're

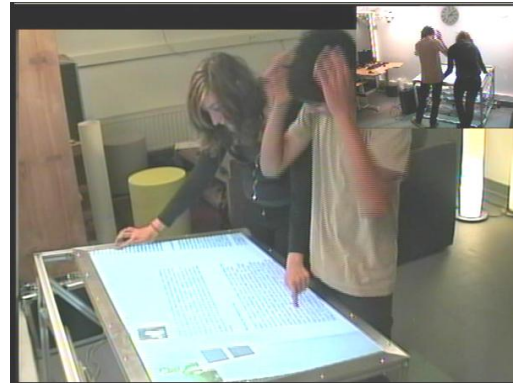using too much space, use your own half!', etc.[1] (see figure 5).

Figure 5: A high-fidelity test participant grabbing his head in frustration, when the other participant covers his work space.

This 'territoriality' behaviour was, however, not related to the use of one common central, circular menu (see figure 3), or the use of individual personal menus per user. Both groups using only one common menu, and groups using more (individual) menus made remarks related to their 'own' side of the multi-touch table.

The territoriality issue is further complicated by users physically switching positions at the table. Often, when one user took over a task (such as typing) from another user, they did not move the applications on the table. Instead, they switched positions at the table, passing behind each other to reach the other part of the table in which, for instance, the typewriter was located. On the other hand, content such as pictures and books were resized and passed on, moving the digital content around instead of physically switching places.

In terms of effort, it is easier to move the functionality (which takes only one gesture), rather than to walk to the other side of the table. The position-switching behaviour suggests that territoriality doesn't mean 'the side of the table with *my content/functionality* on it', as much as 'the side of the table I am *currently working on*'. While working on the collaborative assignment, users switched roles, and switched between working on different kinds of content – after switching, the 'new' content or functionality users were working on quickly became *their* content or functionality. Furthermore, the 'static' functionality and the 'moveable' content suggest that table layout (see the section on table layout) has an important influence on territoriality, as well. Many participants seemed to consider the functionality on the table as part of the more static table layout, while content such as pictures etc. were not

---

[1] Similar behavior has also been reported in [13]. In that experiment, however, users often didn't know other users interacting at the same multi-touch surface. Therefore, their reactions and comments often were more polite or covert.

considered to be part of that static layout. Therefore, functionality was not moved and passed on to other users.

Territoriality: Comparison
Although it was first thought that in the low-fidelity prototype, the typewriter was not moved around due to its weight, a similar observation was made in the high-fidelity tests. Although test users did exchange content (photos, articles,…), tools such as the typewriter tool or the content browser most often were left in its position, with the test users exchanging places.

In addition, low-fidelity and high-fidelity prototypes results were quite different regarding territoriality conflicts: while in the low-fidelity setting, participants collaborated without any problems, the high-fidelity setting created a lot more territoriality issues. When comparing these territoriality issues to table size ratings (see graph 1), it becomes apparent that the type of prototype has an important effect on territoriality. While one would expect more problems to occur when participants judge the work surface as too small (i.e., in the low-fidelity prototype), the table size ratings given by participants show the opposite. In the low-fidelity setting, with a table that was rated as quite limited in work space, no territoriality issues ensued. In the high-fidelity setting, with a table that appeared large enough to most participants, several territoriality issues occurred. As in the participant interaction and table layout sections, it seems that the 2D high-fidelity vs. the 3D low-fidelity work space is an important factor here. Participants get in each other's way more easily on a 2D surface than in a 3D space.

To summarize, whereas the low-fidelity prototype allowed people to divide the available space in a natural way, as described also in the low-fidelity setting of [19], the high-fidelity prototype tests in our study showed different results. Participants ran into territoriality conflicts and, depending on the nature of the items displayed (content vs. applications), they switched between passing on the digital content and physically switching places themselves to leave the table layout intact. This leads to the conclusion that specific issues, such as the 2D vs. 3D collaboration mentioned in the table layout comparison section, and a perceived difference between 'static' functionality and 'moveable' content (based on table layout), can have a profound influence on the way people experience territoriality in the real-world vs. in digital collaboration.

Unused Functionality

Unused Low-Fidelity Functionality
Although the low-fidelity tabletop prototype had a waste bin in which test participants could leave unused article scraps and other content material they did not use, it was not used as it took up too much table space. In the majority of user groups, unused scraps eventually made up the bottom layer in the work zone.

Participants tended to leave all material (both tools and content material) lying around on the table. Even tools that

were used only rarely were not returned, but stayed on the table, ready for use. Participants preferred to have all material on a rather crowded table, rather than use a cleaner, more organised table and go through the trouble of returning the materials and asking them back later.

Unused High-Fidelity Functionality
Like in the low-fidelity tests, test participants didn't use the recycle bin very often in the high-fidelity tests either. Instead, they bended the 'rules' implicitly imposed by the platform, and used other, creative methods to dispose of unused materials. Participants resized the content until it was very small (and nearly invisible), or simply moved the unwanted content off-screen.

As already mentioned in the table layout section, some groups tended to leave all material 'lying around' on the table, like the low-fidelity groups. However, most of them soon got frustrated with the mess of content displayed on the table.

Unused functionality: Comparison
Low-fidelity and high-fidelity tests were very similar in terms of functionality that was left unused by the participants: few test groups used the recycle bin. This suggests that low-fidelity prototypes potentially are adequate predictors of the perceived usefulness of functionality. The prototypes can predict which functionality users think of as useful, and which parts they think are redundant. However, it can be inferred that this predictive power will only hold when the interaction fidelity of the low-fidelity prototype is quite high on a more detailed level, as well. In the case of the recycle bin, lower interaction fidelity probably would have made the bin more useful. It can be assumed that the low-fidelity bin would have been used more often if users were allowed to place it underneath the table, or next to it, like a normal, real-world bin. Increasing prototype fidelity by only allowing the bin to be placed *on* the table made the bin less useful: it took up too much place on the table, compared to the actual benefit users had from having it on the table. Similarly, in the high-fidelity prototype, the bin took up too much screen estate, given the limited amount of functionality it added.

In sum, our findings suggest that, given the fact that interaction fidelity is high at a general level *and* at a lower, more detailed interaction level, low-fidelity multi-touch prototypes can be accurate predictors of the user's perceived usefulness of parts of the prototype.

DISCUSSION

The results discussed above show that there are aspects to low-fidelity multi-touch prototyping that allow for prototyping, and aspects that resist prototyping. The way participants passed on content, but left functionality in place in the high-fidelity setting was similar to the low-fidelity setting, not because of the intrinsic properties of the functionality itself (as in the low-fidelity setting: the weight of the typewriter, etc), but because of the participant's

implicit agreement on table layout. Similarly, the recycle bin was not used in both test series due to a comparable perception concerning its usefulness and the valuable screen estate it demanded.

Table Layout	
Lo-fi	Hi-fi
Available space is divided in personal and common zones	
Organic, fluent	Structured or messy
3D layering	2D overlapping content
Participant Interaction	
Lo-fi	Hi-fi
Group or parallel individual	
No turn taking	Turn taking
3D individual work	2D 'interference'
Territoriality	
Lo-fi	Hi-fi
'Static' functionality and 'moveable' content	
Organic, fluent (no issues)	Territoriality issues
3D static content	2D resizable content
Unused Functionality	
Lo-fi	Hi-fi
No use of the recycle bin	
Leave unused material lying around	Try out creative methods

Figure 6: Similarities and dissimilarities between low-fidelity and high-fidelity prototypes.

However, the possibility to interact in a 3D space vs. on a 2D surface influenced the participant's interactions profoundly. While it was expected that the 3D tangible objects vs. 2D touch surface would influence more low-level interactions, the results show that higher-level interactions with the work surface and between participants are also influenced. This influence on higher-level interactions can be seen as a *side effect* of more obvious, low-level interactions, also affecting the usage of the prototype as a whole [7]. An example of this is the high-level table layout being influenced by the (low-level interaction) possibility to stack items, and create 'layers' of content. Although the influence of 2D vs. 3D on higher-level interactions is often a more indirect influence, its effects cannot be ignored.

We can conclude that when using low-fidelity, tangible prototyping as an evaluation method of high-level interactions, one should be cautious in interpreting the results. Although several low-fidelity observations were similar in the high-fidelity situation, the difference between 2D vs. 3D interaction also influences interaction with the surface and other users at a high level. Due to this influence, similarities between low-fidelity and high-fidelity prototypes stop, and observations start to diverge. It is, however, difficult and unpredictable to pinpoint the exact point at which the observations diverge. As shown in figure 6, similarities between the low and high fidelity tests always lie on a high level of interaction. The dissimilarities found can mostly be related to the prototyping medium used and the specific constraints linked to that medium. For example, paper is a medium that enables three-dimensional

interactions due to its physical presence on a table. This three-dimensional interaction allows users to easily 'layer' paper documents, and retrieve them by picking up the paper. On a 2D surface, on the other hand, one is limited to overlapping content.

Intrinsically linked to the above, the low-fidelity prototyping method presented in this paper is an excellent method for idea generation. In this way, the prototype has the function of a 'filter', as proposed by Lim et al. [7]: it makes it possible 'to examine an idea's qualities without building a copy of the final design'. In the case presented in this paper, the low-fidelity prototype presented ideas about how to shape the interaction. The low-fidelity prototype inspired new concepts such as content layering and having all material immediately available, without 'closing' functionality or using a recycle bin. This type of design ideas did not emerge during high-fidelity testing.

LIMITATIONS AND FUTURE WORK
The study described above highlights some aspects of high and low-fidelity prototyping that are quite universal in multi-touch applications (e.g. the 2D vs. 3D distinction) and have been described in existing literature (e.g. the territoriality issues in [19]). However, future studies will have to confirm if the specific observations described in this paper can be generalised to multi-touch applications in general.

In addition, low-fidelity prototyping has some often-used, adequate methods such as paper prototyping for GUIs, using paper and post-its to create drop-down menus and other controls. For multi-touch tabletops, however, such often-used methods have not yet been established. It would be interesting to see in future work which innovative low-fidelity methods can be developed to make low-fidelity observations more predictive of high-fidelity results.

CONCLUSION
We have reported on a study in which we compare high-level interactions on a physical and digital prototype of a tabletop application. Although the low-fidelity prototype was designed to have a focus on global, high-level interaction, and not on detailed manipulation of objects, test results showed that differences in more detailed interaction and manipulation of objects also have a profound influence on higher-level interactions. In order to create testable low-fidelity multi-touch prototypes, we found that only very high-level user interactions can be generalised towards a high-fidelity prototype. In addition, the low-fidelity prototype approach presented proved to be very valuable to generate design ideas concerning both high and low-level user interactions.

REFERENCES

1. Al Mahmud, A., Mubin, O., Shahid, S., and Martens, J. Designing and evaluating the tabletop game experience for senior citizens. In *Proc. NordiCHI 2008*, ACM Press (2008), 403-406.

2. Apted, T., Kay, J., and Quigley, A. Tabletop sharing of digital photographs for the elderly. In *Proc. CHI 2006,* ACM Press (2006), 781-790.

3. Arnowitz, J., Arent M., and Berger, N. *Effective Prototyping for Software Makers*. Morgan Kaufmann, San Francisco, CA, 2007.

4. Gutwin, C., and Greenberg, S. A Descriptive Framework of Workspace Awareness for Real-Time Groupware. *Comput. Supported Coop. Work 11, 3* (Nov. 2002), 411-446.

5. Kristensson, P. O., Arnell, O., Björk, A., Dahlbäck, N., Pennerup, J., Prytz, E., Wikman, J., and Åström, N. InfoTouch: an explorative multi-touch visualization interface for tagged photo collections. In *Proc. NordiCHI 2008.* ACM Press (2008), 491-494.

6. Lim, Y., Pangam, A., Periyasami, S., and Aneja, S. Comparative analysis of high- and low-fidelity prototypes for more valid usability evaluations of mobile devices. In *Proc. NordiCHI 2006.* ACM Press (2006), 291–300.

7. Lim, Y., Stolterman, E., Tenenberg, J. The anatomy of prototypes: Prototypes as filters, prototypes as manifestations of design ideas. *ACM Transactions on Computer-Human Interaction. 15, 2* (2008), 1-27.

8. Marco, J., Cerezo, E., Baldassarri, S., Mazzone, E., and Read, J. C. Bringing tabletop technologies to kindergarten children. In *Proc. British Computer Society conference on HCI 2009.* British Computer Society (2009), 103-111.

9. Mazalek, A., Winegarden, C., Al-Haddad, T., Robinson, S. J., and Wu, C. Architales: physical/digital co-design of an interactive story table. In *Proc.TEI 2009,* ACM Press (2009), 241-248.

10. McCurdy, M., Connors, C., Pyrzak, G., Kanefsky, B., and Vera, A. Breaking the fidelity barrier: an examination of our current characterization of prototypes and an example of a mixed-fidelity success. In *Proc. CHI 2006,* ACM Press (2006), 1233-1242.

11. Microsoft Surface. www.surface.com.

12. North, C., Dwyer, T., Lee, B., Fisher, D., Isenberg, P., Robertson, G., and Inkpen, K. Understanding Multi-touch Manipulation for Surface Computing. In *Proc. IFIP TC 13 Conf. on Human-Computer interaction: Part II*. Lecture Notes In Computer Science, vol. 5727. Springer-Verlag, Berlin, Heidelberg, 236-249.

13. Peltonen, P., E. Kurvinen, A. Solovaara, G. Jacucci, T. Ilmonen, J. Evans, A. Oulasvirta, and P. Saarikko. It's Mine, Don't Touch!: interactions at a large multi-touch display in a city centre. In *Proc. CHI 2008,* ACM Press (2008), 1285-1294.

14. Piper A.M., Hollan J.D., Tabletop Displays for Small Group Study: Affordances of Paper and Digital Materials. In *Proc. CHI 2009,* ACM Press (2009), 1227-1236.

15. Piper, A. M., O'Brien, E., Morris, M. R., and Winograd, T. SIDES: a cooperative tabletop computer game for social skills development. In *Proc. CSCW '06*, ACM Press (2006), 1-10.

16. Quigley, A., Subramanian, S, and Izadi, S. Special issue on interaction with coupled and public displays. *Personal and Ubiquitous Computing, 13*, 8, Springer-Verlag, November 2009.

17. Saffer, D. *Designing Gestural Interfaces*. O'Reilly, Sebastopol, CA, 2009.

18. Scott, S.D., Grant, K.D., Mandryk, R.L. System Guidelines for Co-located, Collaborative Work on a Tabletop Display. In *Proc. ECSCW 2003*, Kluwer Academic Publishers (2003), 159-178.

19. Scott, S.D., Carpendale, M.S.T. & Inkpen, K.M. Territoriality in Collaborative Tabletop Workspaces. In *Proc. CSCW'04,* ACM Press (2004), 294-303.

20. Sefelin, R., Tscheligi, M., and Giller, V. Paper prototyping - what is it good for? A comparison of paper- and computer-based low-fidelity prototyping. In *Proc. CHI 2003,Extended Abstracts on Human Factors in Computing Systems,* ACM Press (2003), 778-779.

21. Sharp, H., Rogers, Y., Preece, J. *Interaction Design. Beyond Human-Computer Interaction. (2nd. edition)* Wiley, West Sussex, England, 2007.

22. Shen, C., Lesh, N. B., Vernier, F., Forlines, C., and Frost, J. Sharing and building digital group histories. In *Proc. CSCW 2002*, ACM Press (2002), 324-333.

23. Snyder, C. *Paper Prototyping. The Fast and Easy Way to Design and Refine User Interfaces*. Morgan Kaufmann, San Francisco, CA, 2003.

24. Terrenghi, L., Kirk, D., Sellen, A., and Izadi, S. Affordances for manipulation of physical versus digital media on interactive surfaces. In *Proc. CHI 2007*. ACM Press (2007), 1157-1166.

25. Wobbrock, J. O., Morris, M. R., and Wilson, A. D. User-defined gestures for surface computing. In *Proc. CHI 2009,* ACM Press (2009), 1083-1092.

26. Yee, W. Potential Limitations of Multi-touch Gesture Vocabulary: Differentiation, Adoption, Fatigue. In *Proc. HCI International 2009*. J. A. Jacko, (Ed.) Lecture Notes In Computer Science, vol. 5611. Springer-Verlag, Berlin, Heidelberg, 291-300.