

Simple Nonvisual Interaction on Touch Tablets

Sonja Rümelin, Valerie Kroner, Andreas Butz

Human-Computer Interaction Group, University of Munich (LMU)

Abstract

In this paper, we compare different designs for a touch tablet interface to support a simple, repeated task with high visual load in itself. A user-centered approach was applied throughout the development. Expert users were involved in the analysis, design and evaluation of an application for time tracking in a production planning environment. In a lab user study, touch gestures and multi-touch input outperformed screen buttons in terms of visual demand, while error rate and efficiency remained stable. A subsequent expert evaluation in the wild confirmed these results. This shows that for repeated tasks with high visual load, more complex touch interaction may be preferred than simpler but widespread techniques.

1 Simple Interaction

In many everyday situations, we do not need complex and powerful interaction with the devices around us, but would rather prefer simple and non-distracting input. For example, when controlling a music player while walking or driving, we might well accept a small set of commands, such as *play*, *pause*, *skip* and *volume*, if they in exchange can be issued blindly and let us keep our eyes on the road. We studied a similar situation in an industrial environment where production planners need to track work times and rate work quality in the subsequent steps of an assembly line. Their goal is to achieve a better overall distribution of the available time and thereby improve overall quality. The task can be characterized as particularly time-constrained with high visual demand.

In such an assembly line, each worker is assigned a specific *task*, which is performed in *iterations* and in turn consists of *steps*. Production planners capture the times needed for each *step* and the quality of work in order to adjust the scheduled times to fit the workers' capabilities better. In our case, production planners still used a stopwatch, pen and paper to record the single *steps*, which took about 10 seconds each. For each step, the quality was then rated on a scale from 80% to 120%. The results are analyzed to possibly improve the *iteration*. In the course of making this entire process digital, the interface had to be designed

to best fit such a repeated simple interaction. The main focus of the users should lie on the process they observe, while interaction has to provide as little distraction as possible.

2 Initial Analysis

Throughout the development, we followed a user-centered approach. We first interviewed a time-tracking manager to understand the process and to get background information. Then we conducted semi-structured interviews of about two hours with three time recorders to get a detailed insight into the context and requirements from the end users. Two of them, our *expert users*, were later involved in iterative testing. Finally, we observed time recorders in their working environment while they recorded a production process.

A main problem found was the non-ergonomic use of stopwatch, pen and paper, and the time-consuming transfer to an excel sheet. Other findings were the noisiness of the environment and the fact that users cannot put the device down, but have to stand wherever they get the best view onto the process. Most prominently, we found that an easy-to-use interface is needed to perform the quick repeated inputs of time and quality. The device should be held ergonomically, allow quick input, and give users time to watch what they are about to review. In addition, the company favored the use of consumer hardware, which is comparatively cheap and can also serve other purposes.

3 Interface Design

Before starting the actual design, we considered the input features of a tablet and the functionality needed.

3.1 Input Design Space

Multi-Touch - Most current touch recognition hardware can distinguish several touch points. Touching with one, two or more fingers may initiate different actions, as known from devices such as the iPad. The concurrency of multi-touch input can accelerate input, which, in return, might cause an increase of erroneous input (Banovic et al. 2011). Another potential drawback is that missing expectations may result in a higher learning effort for users (Freeman et al. 2009).

Multi-Tap - A common concept from mouse interaction is using single and double (rarely even triple) clicks to trigger different actions. Similar, quick succeeding taps can be used for different commands. Sánchez and Aguayo translated the concept of multi-tap for text input of mobile phones with hard keys to smartphones with touch screens, to enable quick access to multiple characters at one position of the screen. This can be used to increase overall button size (Sánchez and Aguayo 2007).

Touch Duration - The duration of a touch can be used to encode certain values, but calibration and repeatability remains a problem, unless proper feedback is provided.

Touch Gestures - A common example for touch gestures is to swipe left for scrolling pages in an e-book. Adding just other movement directions already creates a set of simple and well distinguishable interactions. Using a restricted gesture set, interaction has shown to require less visual attention than direct touch buttons, as it can be performed on an arbitrary position of a touch screen (Bach et al. 2008, Ecker et al. 2009).

Touch Pressure - Force-sensing surfaces can differentiate between light and strong touch actions. This allows enhancing touch interaction with a further dimension (Richter et al. 2010); however, devices need to be equipped with appropriate sensors.

Text Recognition - Tablets can be used to input text commands without restrictions, such as *Start* or numerical values. However, robust recognition of text is still critical and extra difficulty lies in continuously written text (Sas & Markowska-Kaczmar 2012). Moreover, the temporal performance restricts the application.

Orientation - Many consumer devices contain acceleration sensors to detect device orientation. Van Tonder used tilt for interacting with a map application and could show its potential for an explorational navigation task (Van Tonder & Wesson 2010). However, throughput of information is low (MacKenzie & Teather 2012) especially when recognition becomes unreliable when the device is moving, which makes it unsuitable for our task.

Speech Recognition - Speech recognition allows the input of a large command set, but background noise poses a technical problem with built-in microphones. Without using headsets, the noisy environment prevents its use.

3.2 Required Interaction

For our use case, a limited set of actions is required. Users need to specify the action currently observed as well as quality ratings from 80 to 120 in steps of five (i.e., 9 levels). Action sequences are known beforehand: interaction can be described as a state machine (Figure 1) with at most three actions in each state (Table 1).

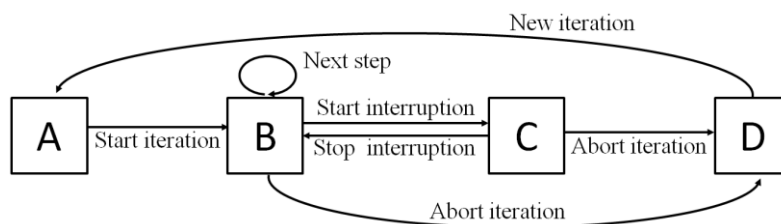


Figure 1. Visualization of state sequences.

In the beginning, only *Start* can be selected (A). When time recording has started, the next step (signaling the end of the current step), an interruption, or the end of recording can be

chosen (B). An interruption can be stopped by starting the next step or aborting the entire recording (C). After the last step, a new iteration can be started or recording can be stopped (D).

State	Action 1	Action 2	Action 3
A	Start	[Interruption]	[Stop]
B	Next	Interruption	Abort
C	Next	[<i>Interruption running</i>]	Abort
D	New iteration	[Interruption]	[Stop]

Table 1. Possible actions during time recording.

Together with the 9 quality levels, twelve input values need to be distinguishable. In addition, an overview of the steps of the current process should be presented for orientation.

3.3 Layout Variants

From the analysis above, we developed five different interface designs (Figure 2), using simple touch and the techniques *multi-touch*, *multi-tap* and *touch gestures*. Two of the layouts use buttons only, either in a vertical or horizontal design. The three large buttons control the possible actions, while nine smaller buttons rate quality. The rest of the screen is used to indicate the current steps.

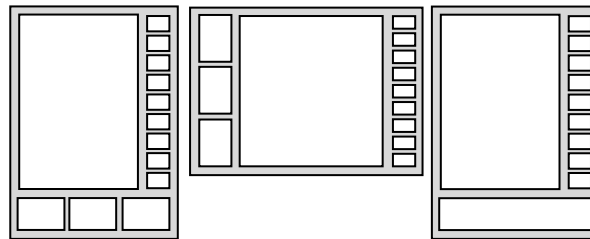


Figure 2. Screen layouts: *PORTRAIT* (left), *HORIZONTAL* (middle), structure of other three layouts (right).

Various different button layouts had been evaluated in an informal pre-study with two expert users. The key insight in this step was that buttons should be placed at the screen borders to support both one-handed and two-handed use. The final button layouts now contain the action buttons on the bottom or left side respectively, and buttons for quality values aligned on the right side. Left-handed people can turn the tablet around to have the layout mirrored.

The other three layouts are designed with the input methods *multi-touch*, *multi-tap* and *touch gesture*. The whole screen is used for input. In the background, the steps as well as the chosen quality value and the explanations of the possible actions are presented (Figure 3). In these last three layouts, quality values (except 100) were not chosen directly, but by

incrementing or decrementing the initial value of 100. We did not use this concept for the button-only layouts because if single touch is used and visual attention is required anyway, directly choosing the required value is obviously much faster. Gestures were designed together with the expert users, and according to established functionality; for example, a swipe from the right, often used to browse forward, was used to start or go to the next step.

In all layouts, input was confirmed by visual feedback. Additional audio or haptic feedback was also given, depending on the level of ambient noise (Hoggan & Brewster 2007).

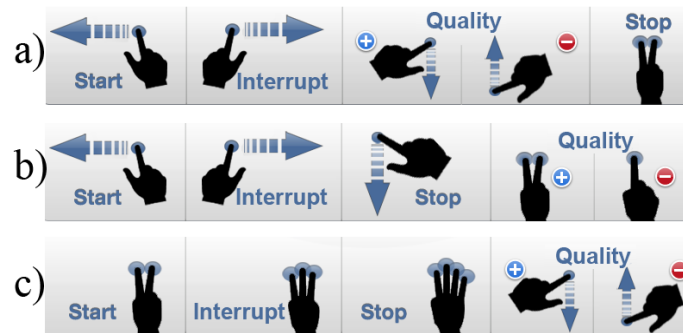


Figure 3. Displayed explanations in the lower display portion for a) SWIPE b) SWIPE-TOUCH c) MULTI-TOUCH.

4 Usability Evaluation in the Lab

A user study was conducted to compare the different designs in a task of repeated simple interaction.

4.1 Participants

Fifteen participants took part in the study (eight male, seven female). They were aged between 20 and 59 (mean = 28), and are not involved in current research. 67% own or have experiences with smartphones. 13% have not used phones or tablets with touch functionality before. One participant was left-handed and used the mirrored layouts.

4.2 Experimental Design

The experiment used a within-subjects design. A pre-study with two usability experts revealed that one of the designs, MULTI-TOUCH, which uses taps with two, three and four fingers, causes a much higher workload and makes commands hard to discriminate (Wobbrock et al. 2009), so it was kept out of the usability evaluation. Therefore, the independent variable *layout variant* contained the four levels: HORIZONTAL, PORTRAIT, SWIPE and SWIPE-TOUCH. The dependent variables were *visual distraction*, *error rate*, *efficiency*, *subjective workload* and *usability*.

4.3 Study Preparation

We could not conduct the user study in the real working environment due to access regulations, so the situation of reviewing a process at the assembly line was reproduced in the lab, which also provided more controlled test conditions. The process of assembling a Fischer Technik™ car was chosen because it can be split up into 19 clearly distinguishable steps and be completed within a maximum of 69.9 seconds (original iteration time). It was performed and filmed four times in advance (Figure 4). Quality values were added so that the participants knew the proper input for quality values. The different layouts were realized in an app. An iPad 2 running iOS5 was used for the study.



Figure 4. Screenshots of the process video. Quality values as well as interruptions are indicated.

4.4 Study Procedure

Participants performed the study individually in a laboratory setting. After an initial introduction of the context and the process to be observed, they were presented with the first layout, could explore it and ask questions about the functionality. They captured two iterations with this system variant and then filled out the NASA Task Load Index (Hart & Staveland 1988) and USE questionnaire (Lund 2001). The sequence of layouts was counterbalanced using a Latin square. After the last test, participants ranked all variants regarding the USE criteria and were asked further comparison questions. They were encouraged to think aloud and give feedback at all points during the study. All study sessions were videotaped for analysis and manual gaze tracking afterwards.

4.5 Results and Discussion

In the following, we will describe the results from our quantitative and qualitative analysis. If not otherwise stated, they are reported at a significance level of .05.

4.5.1 Visual Distraction

Visual distraction from the main observation task was measured by counting the number of glances on the screen performed mainly to control input (see Figure 5). A one-way ANOVA showed a significant effect of the tested layout ($F_{3,56} = 22.01$).

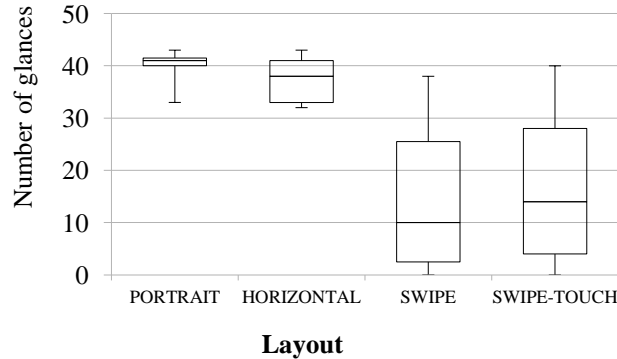


Figure 5. Visual distraction of the different layouts.

With an effect size of $\epsilon > 0.5$, the impact was found to be large, meaning that the layout significantly influenced visual distraction. A post-hoc Duncan's multiple range test revealed two homogeneous groups: PORTRAIT and HORIZONTAL require more control glances compared to SWIPE and SWIPE-TOUCH. Within these groups, no significant differences could be found. A possible explanation is the higher difficulty of hitting a button than performing a gesture of multi-touch input correctly. Participants were also asked about their subjective impression of visual distraction. Their answers confirm the results and indicate that frequent glancing correlates with perceived distraction.

Error Rate & Efficiency

The error rate was measured as wrong, uncorrected inputs, while efficiency was measured by the deviation from the optimal input. The more corrections are made, the lower the efficiency. SWIPE and SWIPE-TOUCH required more overall input than the button-based layouts, because a quality value required several interactions ($90 = 100 - 5 - 5$). Overall, all layouts had a low number of errors and corrections (Table 2). We report no significant layout effect.

		POR	HOR	SWI	ST
Errors	Min/Max	0/2	0/1	0/2	0/1
	Median	0.40	0.33	0.53	0.60
Optimal inputs	Taps/ Swipes	73/0	73/0	0/106	38/59
Additional inputs	Min/Max	1/7	0/6	0/8	0/8
	Median	2.60	1.67	2.20	2.67

Table 2. Errors, optimal and additional inputs.

4.5.2 Subjective Workload

An analysis of the perceived task load captured using the NASA Task Load Index showed a low task load for all layouts (41-48 of 120). There was no significant difference between the various layouts. SWIPE and HORIZONTAL achieved the best results.

4.5.3 Usability

The USE questionnaire was used to measure usability in terms of four categories. By trend, SWIPE always achieved the best rating. However, the only significant difference was revealed using a Kruskal-Wallis test ($F_{3,60} = 13.64$) regarding *ease of learning*. Ranking all layouts regarding the USE categories confirmed the results. SWIPE was rated best in all categories except the ease of use category, in which HORIZONTAL performed best. SWIPE-TOUCH was rated worst except for usefulness, in which PORTRAIT was regarded to perform worse. Comments during the study also support the good results of SWIPE. “*I did not need to concentrate too much on the device*” and “*I felt less mentally demanded when using gestures than when hitting a button*”.

4.6 Comments and Further Observations

No correlation between experience with touch screens and performance could be found in any category. Regarding the holding position of the device, participants were mostly holding it with two hands and interacting with their thumbs when using PORTRAIT (73%) and HORIZONTAL (100%). They were undecided with SWIPE and SWIPE-TOUCH where in both layouts, 54% held the device with two and 46% with one hand.

5 Expert Evaluation

To substantiate the results of the lab study, we conducted a follow-up test with two expert users who frequently perform time tracking. They tested all four layouts in the real working environment, tracking real times along the assembly line. The study procedure was similar, except that they recorded six cycles with each layout instead of two.

The expert users rated SWIPE and SWIPE-TOUCH to produce the lowest overall workload, captured by the Nasa TLX questionnaire, with mean ratings of 22 and 26 (of 120), whereas the button layouts PORTRAIT and HORIZONTAL were rated with 33 and 38, respectively. Compared to values taken in the analysis phase in the beginning, where the overall workload of manual capturing was rated with an average value of 70, the electronic variants all proved to be less demanding. Although the new input method via touch tablet was just introduced, workload could be reduced by 50%.

Regarding usability, participants stated that SWIPE and SWIPE-TOUCH perform better than the button layouts, especially in terms of usefulness and satisfaction (see Figure 6). Even ease of learning was not considered worse. Overall, the gesture-based layouts were preferred even more by the expert users than by the participants of the lab study. They emphasized that

they “*help to concentrate on the process, not on the device*”. Asked to give an overall ranking, they agreed on the following order: SWIPE, SWIPE-TOUCH, HORIZONTAL and PORTRAIT.

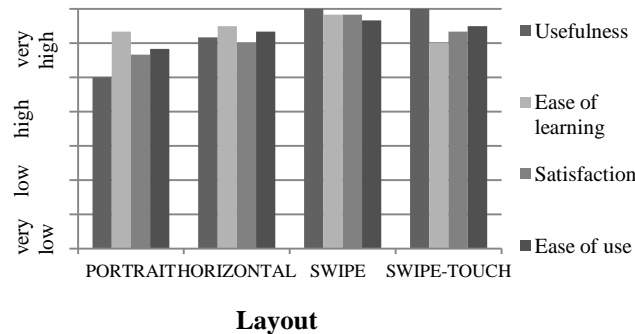


Figure 6. Usability ratings of expert users.

6 Conclusion

In this paper, we compared different approaches for controlling simple functionality on a touch tablet. Our main focus was to design an interface that requires the least possible visual load while holding workload and error rate at a low level.

In a preliminary expert evaluation, our set of multi-touch input containing two-, three- and four-finger touches at a time was regarded as too complex. We compared two button layouts against two layouts using multi-touch and touch gestures. In a lab study, those complex layouts showed a significantly lower visual demand, while efficiency was still as high as with the button layouts. An expert evaluation in a real working environment showed that workload was lower with the layouts using multi-touch and touch gesture. Usability was also rated better. The results of this study can inspire the design of interfaces for simple, repeated tasks during high visual workload. Multi-touch and touch gestures may require a certain learning phase. However, after that phase, they can outperform simple touch input in terms of visual distraction while efficiency remains high and perceived workload is kept low.

One requirement was to use multipurpose consumer hardware, which provided several restrictions to the input and output. To better fit the process of time tracking, we will explore a specially designed device, such as a stopwatch with extended functionality (e.g., data export).

References

- Bach, K. M., Jæger, M. G., Skov, M. B., & Thomassen, N. G. (2008). You Can Touch, but You Can't Look: Interacting with In-Vehicle Systems 1. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, ACM Press, 1139–1148.

- Banovic, N., Li, F., & Dearman, D. (2011). Design of unimanual multi-finger pie menu interaction. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS)*, ACM Press, 120-129.
- Ecker, R., Broy, V., Butz, A., & De Luca, A. (2009). pieTouch: A Direct Touch Gesture Interface for Interacting with In-Vehicle Information Systems. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI)*, ACM Press, 1–10.
- Freeman, D., Benko, H., Morris, M. R., & Wigdor, D. (2009). ShadowGuides: visualizations for in-situ learning of multi-touch and whole-hand gestures. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS)*, ACM Press, 165–172.
- Hart, S. & Staveland, L. (1988) Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock and N. Meshkati, eds., *Human mental workload*. Amsterdam: North Holland Press, 139–183.
- Hoggan, E. & Brewster, S. (2007). Designing audio and tactile crossmodal icons for mobile devices. In *Proceedings of the 9th international conference on Multimodal interfaces (ICMI)*, ACM Press, 162-169.
- Lund, A.M. (2001) Measuring Usability with the USE Questionnaire. *Usability Interface 8*, 2, http://www.stcsig.org/usability/newsletter/0110_measuring_with_use.html.
- MacKenzie, I., & Teather, R. (2012). FittsTilt: the application of Fitts' law to tilt-based interaction. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction (NordCHI)*, ACM Press, 568–577.
- Richter, H., Ecker, R., Deisler, C., & Butz, A. (2010) HapTouch and the 2+ 1 State Model: Potentials of Haptic Feedback on Touch Based In-Vehicle Information Systems. In *Proceedings of the Second International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, ACM Press, 72–79.
- Sánchez, J., & Aguayo, F. (2007). Mobile messenger for the blind. In S. Constantine & M. Pieper (Eds.), *Universal Access in Ambient Intelligence Environments*. Heidelberg: Springer, 369-385.
- Sas, J. & Markowska-Kaczmar, U. (2012). Similarity-based training set acquisition for continuous handwriting recognition. *Information Sciences* 191, May, 226-244.
- Van Tonder, B. & Wesson, J. (2010). Is tilt interaction better than keypad interaction for mobile map-based applications? In *Proceedings of the 2010 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists (SAICSIT)*, ACM Press, 322-331.
- Wobbrock, J.O., Morris, M.R., & Wilson, A.D. (2009) User-Defined Gestures for Surface Computing. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, ACM Press, 1083–1092.

Contact Information

sonja.ruemelin@ifi.lmu.de - <http://www.medien.ifi.lmu.de>