User-Acceptance of Latency in Touch Interactions

Walter Ritter[(✉)], Guido Kempter, and Tobias Werner

Vorarlberg University of Applied Sciences, Dornbirn, Austria
`walter.ritter@fhv.at`

Abstract. Nowadays direct input interfaces are nearly ubiquitous due to the advent of touch screen based smartphones, tablets, and computers. Latency in human-computer interfaces has been discussed since a long time, but established numbers have been questioned in recent research regarding their applicability for direct input interfaces like touch screens. This pilot study focuses on user-acceptable levels of latency in two distinct tasks: simple tapping on interface elements for invoking an action, and dragging tasks to control analog settings. Our results show acceptable latency levels around 300 ms for low attention tapping tasks and around 170 ms for dragging tasks, where visual feedback is essential. These findings are in accordance to previous findings and confirm the importance of considering the task to be fulfilled for drawing conclusions.

Keywords: Latency · Touch · Interaction · Acceptance · Gesture recognition

1 Introduction

Response times of technical systems have been a subject of debate since a long time [6]. Since the success and ubiquity of direct input devices like touch screens for smartphones, tablets, and computers, the question has been raised if findings for indirect input interfaces still apply (see [4,8]).

From a technical perspective, longer accepted response times would allow for example for more freedom regarding system design (e.g. network based architectures), or more reliable recognition of gestures. One recent example for this is the delay of 300 ms present in webviews of current smart phones and tablets to distinguish between a tap on a link and a double tap gesture to automatically zoom webpages to display the content tapped[1]. From a user's perspective, however, shorter response times allow for more efficient use of interfaces. Systems with longer response times might be considered less attractive than a more responsive one, or even prevent the user from successfully fulfilling tasks [12].

---

[1] A discussion of the 300 ms delay in webviews and efforts to work around it can be found at http://blogs.telerik.com/appbuilder/posts/13-11-21/what-exactly-is.....-the-300ms-click-delay.

Miller described these two aspects as technical needs and psychological needs for system response times [6]. Both need to be taken into account when designing systems.

In our research we use a definition of response time (latency) as the time between the moment a user performs an input action and the moment feedback is given by the system (so it's a combination of input-, processing- and output latency). In this paper we only consider visual feedback.

2   Related Work

Already in 1968 Miller pointed out that the needs for system response times vary heavily between different classes of human actions and outlined 17 different scenarios [6]. He proposed that longer delays may be accepted after closure of an activity than what would be accepted during an ongoing activity (clump), due to limitations of short term memory. Distractions for the short term memory would become even more of a problem, if an individual had an awareness of waiting, which usually happens after around 2 seconds according to Miller.

Shneiderman adds that expectations regarding response times are influenced by prior experiences made by users. If users can complete their tasks quicker than before they will be pleasantly surprised. However, if it would be too quick, they might be worried that they didn't perform the task correctly, or if it would be taking much longer they might become frustrated [9]. He also points out that time expectations vary greatly among individuals and across tasks, and that such expectations of people are highly adaptive. Therefore what used to be acceptable a few years ago, might now be considered unacceptable.

Even a small variation in response times might have effects on the perception of delays. Miller noted that 75 percent of test subjects recognized a variation of 8 percent in delays for durations of 2s to 4s [6]. Gallaway proposed a maximum variation of plus/minus 5 % for response times of up to 2s [3]. However, Shneiderman suggests that modest variations up to plus/minus 50 % are still tolerable [10].

Card et al. advocate the use of three different task classes for response times: 100 ms for perceptual processing, 1s for immediate responses, and 10s for unit tasks [2]. Shneiderman mentions 50–150 ms for cursor movements, 1s for frequent simple tasks, 2s to 4s for common tasks and 8s to12s for complex tasks [10].

Indirect input devices like trackpads or mice rely on feedback given on the screen in form of a cursor. This feedback is the only anchor point for a user to verify the correct input actions. In contrast to this, direct input devices like touch-screens would not need input feedback as the user's finger could directly act as zero latency feedback. This might have consequences for the perception of response times. Direct input devices might therefore be more forgiving regarding longer response times since they could simply be ignored.

Jota et al. addressed this question and found that latency in direct input devices still affects interaction. According to them, latency mostly affects movement times during the final stages of pointing. But more importantly in their

research they also observed a significant increase of user performance with decreasing latencies down to 10 ms. They concluded that a reasonable time window to give feedback would be between 20 ms and 40 ms. For comparison, current touch screen devices feature response times between 50 and 200 ms [4].

In a previous study about perceptible levels of latency Ng et al. showed that during dragging operations users were able to perceive latencies down to 2.38 ms [8].

Anderson et al. performed a study about acceptable levels of latency for common tasks with touch screen devices. They reported that delays above 580 ms were considered unacceptable by their users [1]. However, as the experimental task were short, higher delays might be acceptable for more complex tasks.

While some studies confirm that latency can be perceived down to levels way beyond the capabilities of what is currently available on the market, and that those small latencies still affect performance of users, other studies show that user-acceptable levels of latencies are highly depending on the tasks performed by the users and might be well above the latency for cursor-feedback proposed by Shneiderman [10].

The broad range of findings for touch screen interactions, led us to perform a brief pilot study in the context of a concrete usage scenario to judge possible consequences for the technical requirements for the development of a specific device, balancing the technical and psychological needs.

3   Method

To find out about acceptable response times for our usecases in a specified context we conducted two pilot studies with ten test persons each (6 male, 4 female with an average age of 43 years). All of them already had experiences with touch screen devices. In the first study the acceptance level was rated by means of direct delay comparisons, in the second study delays were rated individually. Both studies should evaluate the acceptance levels for a simple tapping tasks as well as for dragging tasks, where visual feedback was required to be able to fulfill the tasks.

3.1   Delay Comparisons

In a first study participants were asked to tap consecutively on a row of buttons and after each tap wait for a glowing lamp feedback to appear on screen before they moved on. Once they completed a row, they were asked up to which button they found the delay acceptable. Then they had to perform a set of dragging tasks, moving a number over a line at the top of the screen (see Fig. 1 for an illustration of the task screens, and Fig. 2 for an illustration of the rating screen).

In this setup 4 sets of delays were used. Each set was used twice - one time ordered ascending, the other time ordered descending. All participants had to rate all 4 sets in both directions. The order in which test persons were presented with the different sets was randomized to compensate for order effects. The sets
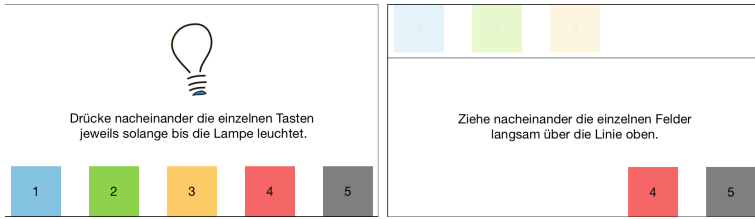
Fig. 1. On the left the user interface for the tapping tasks is shown, on the right side the one for the dragging tasks including the instructions for the test participants.

were comprised of delays between 70 ms (the native latency of the test device) up to 1000 ms.

```
S1: {70,100,150,200,250} ms
S2: {70,150,250,350,500} ms
S3: {70,200,350,600,1000} ms
S4: {70,300,400,700,800} ms
```

To further investigate the influence of the usage-situation on accepted response times, one half of the test persons were instructed to perform the tasks as quickly as possible, while the other half of the test persons were instructed to perform the tasks slowly. Also these instructions were assigned randomly.

Our hypotheses were that we should find differences in the acceptable delays between tapping and dragging tasks (H1a) and find differences in the acceptable delays between the groups with quick and slow instructions (H1b). This would confirm the dependence on the task and the situation for acceptable delay levels.

The tests were performed on an iPad Air with a native latency of 70 ms. Each person needed about 10 min to complete the test.
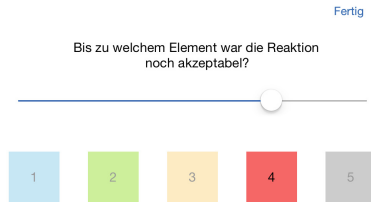
Fig. 2. The rating screen for the last acceptance level of latency. Note that the slider allowed to rate in between two delays too.

3.2   Absolute Delay Ratings

In a second study participants were asked to perform taps on buttons and observe the feedback given for 4 times, and then adjust the hue of one filled rectangle
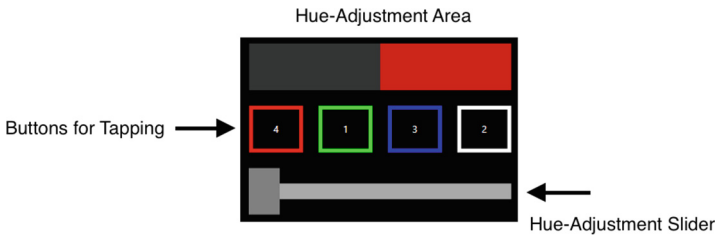
Fig. 3. Tapping and dragging elements for hue adjustments between the left and right rectangle on the top.

to match the hue of another filled rectangle (see Fig. 3 for an illustration of the tasks). This task required exact feedback and was not directly influenced by Fitts law (i.e. there was no direct mapping of the location of the users finger and the feedback given) in contrast to other studies performed (e.g. [4,5,11]).

After each delay time, users were asked to rate the acceptance level of the delay between 0 (best) and 10 (worst) for the tapping task as well as for the dragging task. A complete test consisted of delays of {100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1100} ms where each delay was presented two times, one time in an ascending manner and the other time in a descending manner relative to the previous value. The order itself was randomized to account for order effects.

Our hypotheses was that ratings regarding acceptance levels should differ between the two task groups (H2).

The tests were performed on an Asus EeeTop Touch PC with a native latency of 100 ms. Each person needed about 30 min to complete the test.

4   Results

4.1   Delay Comparisons

To identify an acceptance threshold of latency among the different sets of comparisons we linearly transformed the values of the given ratings into ms values and calculated descriptive statistics. Then we first compared the overall difference between the groups of tap- and dragging tasks using a paired-samples t-test.

To see the effect of the instruction given, we then splitted the results by instruction type and again compared tap- and dragging tasks for differences in the threshold levels using a paired-samples t-test.

The analysis shows that there were differences in the acceptance threshold for latency between tapping and dragging tasks (Tap: avg=263.0 ms, sd=127.6; Drag: avg=212.7 ms, sd=162.8; p=0.002), supporting our H1a despite the high standard deviation for the dragging operation rating.

Comparing the differences regarding acceptance threshold for latency between tapping and dragging tasks considering the instructions given

Table 1. Acceptantce Thresholds for Latency

		Tap		Drag		
Instruction	n	AVG	STDEV	AVG	STDEV	p
Overall	80	263.0	127.6	212.7	162.8	0.002*
Quick	40	286.9	147.5	256.5	206.8	0.234
Slow	40	239.1	100.3	169.0	83.5	0.000*

(Tap: avg=286.9 ms, sd=147.5; Drag: avg=256.5 ms, sd=206.8; p=0.234 for the quick instruction; Tap: avg=239.1 ms, sd=100.3; Drag: avg=169.0 ms, sd=83.5; p=0.000 for the slow instruction) reveals significantly lower acceptable latency levels for slow dragging tasks than for tapping tasks. See Fig. 4 for a graphical representation and Table 1 for a detailed overview of the results. While H1b was not supported for simple tapping tasks, it was confirmed for slow dragging tasks.

4.2   Absolute Delay Ratings

To compare the given ratings for the tested touch latency times we calculated descriptive statistics for each latency and interaction for the interaction types tap and drag.

Results show negative ratings regarding acceptance for tap actions starting at 600 ms and 450 ms for drag actions, however, the confidence intervals overlap considerably in these ranges.

We then compared the two interaction types against each other using paired-samples t-tests to see if there were significant differences in ratings between the two interaction types among the tested latencies.

At latency times of 200 ms, 250 ms and interestingly 400 ms a significant difference between tap and drag interaction was found confirming our H2 for these latencies (see Table 2 for detailed results). Figures 5 and 6 show the ratings for the tested latency times for tap and drag actions.
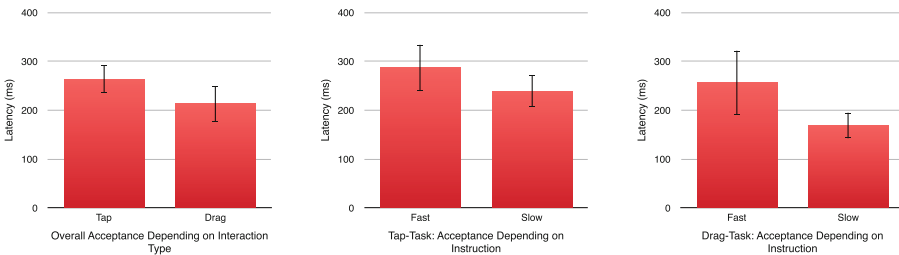
Fig. 4. Differences of acceptance levels depending on interaction type and instruction type. Error-bars show the 95 % confidence interval.

**Table 2.** Latency ratings for Tap and Drag actions

| Tap		Tap		Drag		
Latency (ms)	n	AVG	STDEV	AVG	STDEV	p
100	20	0.6	0.7	1.2	1.8	0.078
200	20	1.3	1.3	2.9	2.7	0.006*
250	20	2.8	1.9	4.2	2.8	0.018*
300	20	2.3	1.7	3.5	2.8	0.079
350	20	4.6	2.4	4.5	3.0	0.875
400	20	3.1	1.7	4.6	3.1	0.027*
450	20	5.6	2.4	5.8	3.4	0.694
500	20	5.4	2.6	6.2	2.6	0.145
600	19	7.3	2.1	7.3	2.7	0.970
700	20	7.7	2.3	7.7	2.5	0.980
800	20	7.4	2.5	8.2	2.4	0.115
900	20	7.8	2.6	8.4	2.2	0.288
1100	20	9.2	1.4	9.2	1.4	0.949

5 Discussion

The results in our study on the one hand confirm findings by Anderson [1] with an absolute negative rating for acceptance starting at 600 ms latency for tap actions, and 450 ms for drag actions, but on the other hand also show the dependency of these levels on the task to be performed, which is inline with previous findings e.g. by [6,9].

The results of our comparison study show the influence of different usage contexts on the level of acceptable latency. In scenarios where not a lot of attention is needed to complete a task (in our case the quick tapping scenario) the accepted threshold is much higher (286.9 ms) than for interactions that involve more attention to detail (in our case the slow dragging scenario) (169 ms). For simple tapping tasks the influence of instructions could not be confirmed.

While our study for absolute delay ratings seemingly yielded much higher acceptable latency times (up to 600 ms) than our comparison study (286.7 ms), a closer look reveals, that the actual level might be in line with what our comparison study revealed. Significant differences between tap and drag tasks in the 200–250 ms range indicate, that the requirements for tap and drag actions are different at this stage. A possible explanation for this would be, that for tap actions these times are already pleasant enough for the users, whereas for drag actions there is still a desire for an improvement. At levels above 300 ms both could be beyond a comfortable levels, and thus not yield significant differences anymore. Interestingly a significant difference was also found at 400 ms. Here it would be interesting to perform additional tests to see if this was due to the
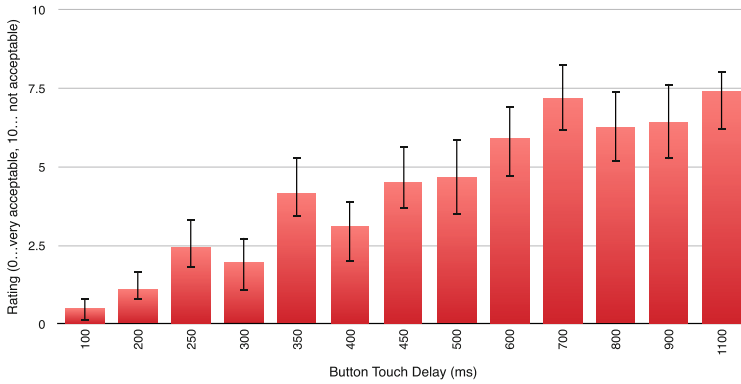
Fig. 5. Differences of acceptance ratings for the tested button touch latencies. Error-bars show the 95 % confidence interval.
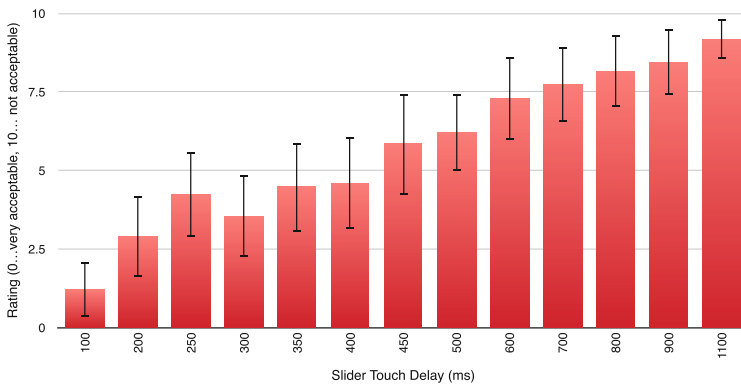
Fig. 6. Differences of acceptance ratings for the tested slider touch latencies. Error-bars show the 95 % confidence interval.

small sample size or if there might even exist something like an uncanny valley [7] for response times.

We are aware that these pilot studies can only be considered as an indication of what might be an acceptable amount of latency in touch interactions due to the small sample sizes. However, since the results are inline with previous research and the numbers are robust with different analysis methods we feel confident that within the narrow usage context we set out in our studies (we worked with very small movements (max. up to 10 cm) specific for our needs), values at around 170 ms for dragging tasks and around 300 ms for low attention tapping tasks are realistic acceptance levels for latency. Other dimensions might yield different requirements.

Shneiderman outlined the importance of prior experience for the judgement of latency [9]. This means that with more responsive technology evolving over the years, these acceptable levels might be subject to change considerably in

future. Jota et al. already confirmed that productivity increases further with even lower latency levels [4], so if technically possible at a reasonable cost, we would recommend aiming for lower latency levels as currently found acceptable by users.

References

1. Anderson, G., Doherty, R., Ganapathy, S.: User perception of touch screen latency. In: Marcus, A. (ed.) HCII 2011 and DUXU 2011, Part I. LNCS, vol. 6769, pp. 195–202. Springer, Heidelberg (2011)
2. Card, S.K., Robertson, G.G., Mackinlay, J.D.: The information visualizer, an information workspace. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 181–186. CHI 1991. ACM, New York, NY, USA (1991)
3. Gallaway, G.R.: Response times to user activities in interactive man/machine computer systems. Proc. Hum. Factors Ergon. Soc. Ann. Meet. **25**, 754–758 (1981)
4. Jota, R., Ng, A., Dietz, P., Wigdor, D.: How fast is fast enough?: a study of the effects of latency in direct-touch pointing tasks. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI 2013, pp. 2291–2300. ACM, New York, NY, USA (2013)
5. MacKenzie, I.S., Ware, C.: Lag as a determinant of human performance in interactive systems. In: Proceedings of the INTERACT 1993 and CHI 1993 Conference on Human Factors in Computing Systems. CHI 1993, pp. 488–493. ACM, New York, NY, USA (1993)
6. Miller, R.B.: Response time in man-computer conversational transactions. In: Proceedings of the Fall Joint Computer Conference on 9–11 December 1968, Part I. AFIPS 1968 (Fall, part I), pp. 267–277. ACM, New York, NY, USA (1968)
7. Mori, M.: The uncanny valley. Energy **7**(4), 33–35 (1970)
8. Ng, A., Lepinski, J., Wigdor, D., Sanders, S., Dietz, P.: Designing for low-latency direct-touch input. In: Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology. UIST 2012, pp. 453–464. ACM, New York, NY, USA (2012)
9. Shneiderman, B.: Response time and display rate in human performance with computers. ACM Comput. Surv. **16**(3), 265–285 (1984)
10. Shneiderman, B., Plaisant, C.: Designing the User Interface: Strategies for Effective Human-Computer Interaction, 4th edn. Pearson Addison Wesley, Boston (2004)
11. Ware, C., Balakrishnan, R.: Reaching for objects in vr displays: lag and frame rate. ACM Trans. Comput.-Hum. Interact. **1**(4), 331–356 (1994)
12. Zhicheng, L., Heer, J.: The effects of interactive latency on exploratory visual analysis. IEEE Trans. Vis. Comput. Graph. (Proceedings InfoVis 2014) **20**(12), 2122–2131 (2014)