Does Practice Make Perfect? Learning to Deal with Latency in Direct-Touch Interaction

Elie Cattan[1], **Amelie Rochet-Capellan**[2], **Pascal Perrier**[2], **Francois Berard**[1]

LIG, University of Grenoble[1]
Grenoble, France
first.last@imag.fr

GIPSA-lab, University of Grenoble[2]
Grenoble, France
first.last@gipsa-lab.grenoble-inp.fr

ABSTRACT

Touch latency has been shown to deteriorate users' performances at levels as low as 25 ms, but this was tested only in short experimental sessions. Real life usage of touchscreens covers much longer periods. It provides training which could lead to reduce the impact of latency.

We investigate users' ability to compensate for touch latency with training. Two groups of participants were trained on a tracking task during ten different days over two weeks with either high or low latency. The gap of performances between the two groups, observed at the beginning of the experiment, was reduced by 54 % after training. Users can thus compensate for latency, at least partially. These results nuance the negative effects of touch latency reported in previous work. They suggest that long-term studies could provide better insights on users' behaviors when dealing with touch latency.

ACM Classification Keywords

H.5.2 User Interfaces: Input devices and strategies (e.g., mouse, touchscreen)

Author Keywords

latency; direct-touch; learning; training; tracking task; user performances

INTRODUCTION

Touch latency, the delay between a user's touch input and the resulting feedback, has been shown to decrease users' performances at levels as low as 25 ms [5, 15]. Current commercial touchscreens exhibit latencies in the range of 50 – 100 ms [8], at which users' performance is still strongly affected. Touch interaction is yet more and more present in our everyday life on devices such as smartphones, tablets or larger interactive surfaces. This recurrent use of touchscreens with substantial levels of latency suggests that people may become used to a digital world that answers to their actions with delay. It is unclear, however, if they also learn to *compensate* for latency over time. In other words, after a prolonged use of lagging touchscreens, are people able to execute tasks at the same performance level as without latency?

The motor control literature provides insights into perturbation compensation in the context of visuomotor recalibration experiences. The human brain was shown to quickly compensate for different kinds of perturbation such as visual prism distortion and virtual force fields [30]. People also compensate for delayed visual feedback [6, 10, 21], but it is unclear how long it takes for this compensation to be efficient. These studies also focused on indirect interaction with the hand hidden from the participant's view. This situation is notably different from the direct touch situation in which both the hand and the delayed feedback are visible.

For direct touch, the HCI literature has recently provided several studies on the impact of touch latency on perception or performance [1, 5, 9, 15, 16, 24, 25]. Participants of these studies, however, are only exposed to latency for a short duration (usually less than an hour). In real life, people use their devices every day, during longer periods of time, repeating some tasks several times a day. Real-life usage, thus, offers much more training and opportunity to develop compensatory strategies to counteract the latency of the device. The detrimental effects of touch latency could then be different after heavy exposure than at first exposure.

In this work, we investigate the following questions:

- Do users compensate for latency with practice? If they do, is the compensation partial or total?
- How long does that take?
- How long does that last?
- Is this compensation limited to a specific task or can it be transferred to similar tasks?

We focus on a tracking task. Two groups of participants are trained on the task during 10 sessions spread on different days over two weeks. The *test* group deals with a latency comparable to that of current commercial devices while the *control* group is trained with an almost unnoticeable latency. Our goal is to evaluate if the test group, whose performances should be negatively impacted by latency at the beginning of the experiment, could catch up with the performances of the control group with training.

In order to better understand changes induced by long-term exposure to latency, we also test for transfer of learning to

CHI 2017, May 6-11, 2017, Denver, CO, USA.

ACM ISBN 978-1-4503-4655-9/17/05 ...$15.00.
http://dx.doi.org/10.1145/3025453.3025585

another trajectory. If the compensation is not specific, it would diminish greatly the negative influence of touch latency since users would be able to deal with the delayed feedback on a second task without a new learning phase. We also test for long-term retention to see how long the compensation last.

The presentation of our work is divided as follow. First, we introduce previous work on the detrimental effect of latency as well as on touch latency compensation and transfer of learning. We then explain the design choices of our user study, which was meant to assess the effect of touch latency after several days of training. Finally, we present and discuss the results of the study and their contribution to the understanding of the touch latency effect on users' behaviors.

RELATED WORK

Detrimental effects of latency on interaction

The effects of latency on users' performances were first studied for mouse interaction. MacKenzie and Ware showed that latency decreases users' performances for pointing tasks and that the effect increases with task difficulty [20]. When latency increased from 8 ms to 225 ms, movement time increased by 63 % and error rate more than tripled. They also observed a floor effect since no difference in performance was found between 8 ms and 25 ms. This result was recently refined in a study by Friston et al.[11]. They evaluated users' performances at various levels of mouse latency on both Fitts' and Steering tasks, and showed that mouse latency begins to affect performances at around 16 ms. The impact of latency with a mouse has also been studied for tracking tasks by Pavlovych and Gutwin and then by Ivkovic et al. who showed that above 50ms, latency significantly impacts the precision of the tracking [14, 26].

In direct touch interaction, the temporal delay translates into a spatial gap between the user's finger and the feedback (a cursor, or a dragged object for example). This makes latency more perceivable than in indirect interaction. Ng et al. and Deber et al. measured the Just Noticeable Difference (JND) that users were able to perceive when dragging a square on a touchscreen. They showed that, on average, users can detect latencies down to 6 ms when dragging a square [25] and that they are capable to perceive a latency improvement of 8.3 ms [9]. However, this perception is strongly dependent from the task. Deber et al. showed that the perception threshold is higher for tapping than for dragging [9]. Ng et al. and Annett et al. showed that tasks implying more cognitive load lead to higher thresholds of perception [1, 24]. Focusing on the effect of latency on *users' performances* rather than on the *JND*. Jota et al. showed that users' target acquisition time in direct touch interaction is affected at latency levels as low as 25 ms [15]. Contrary to mouse interaction, no evidence of a floor effect was found. These results were extended to higher task difficulties by Cattan et al. [5].

While latency has been shown to affect performances at levels as low as 25 ms, commercial touchscreens still exhibit latency values in the range [50-100] ms [8]. Software solutions based on prediction have been proposed to compensate for latency [5, 13, 22]. Users, however, by getting used to endure the latency of their devices, may also learn to develop themselves compensatory strategy by changing their behaviour when facing latency. We did not find any elements to confirm or refute this hypothesis in the HCI literature since all aforementioned studies report brief exposures to high latencies, i.e. less than an hour in a single session. Here, we investigate participants' ability to compensate for latency through an experimental paradigm involving a longer training.

Adaptation to delayed feedback

Adaptation to delayed feedback has been studied for a long time in the motor control literature. We review studies which provide some elements of answer to our questions, and in particular "do people compensate for latency?" and "how long does it take?"

Touch interaction can be divided in tapping actions and dragging actions. Adaptation to tap latency has been studied thoroughly in the motor control literature. Temporal Order Judgment or Simultaneity Judgment tasks enable to compute the Point of Subjective Simultaneity (PSS) between a tap from the user and a visual feedback. The PSS is naturally shifted in negative delay value: the feedback feels simultaneous only when appearing with some delay, a phenomena known as the negative asynchrony [2, 16]. This might partially explain why the latency in tapping interaction is less perceivable than in dragging [9]. When continuously exposed to larger delays, users' PSS shifts even more negatively than the natural negative asynchrony, showing an adaptation to the delay [17, 31]. Keetels and Vroomen observed that the temporal recalibration remains small (between 9.4 ms and 16.1 ms) compared to the experiment delay (100 ms asynchrony) [17]. They make the hypothesis that the PSS shift might become bigger with longer exposures.

Latency for visual feedback in tap interaction is not an important bottleneck in HCI because the latency of current devices is already close to the 69 ms perception threshold [9]. Moreover, people are able to recalibrate their perception when exposed to delayed feedback. We thus chose to focus on touch dragging, where latency appears as a notable hindrance for touch interaction.

Latency compensation for continuous actions, like moving an object or a cursor, has been studied in the motor control literature but for indirect interaction using a mouse or a joystick. Cunningham et al. studied feedback delay compensation using a mouse, doing one dimensional motion [6]. Participants had to play a game, moving a plane horizontally on the screen to avoid obstacles coming up. Participants were tested in a single session experiment with a low delay first and then trained with 235 ms delay. After training, participants were roughly as good as with low delay. This indicates a quick and complete compensation from the subjects. Foulkes and Miall also uncovered latency compensation when tracking a target with a joystick [10]. Three groups of participants with different latency levels (0 ms, 200 ms and 300 ms) were trained over 2-3 days on an unpredictable tracking task. A clear improvement was seen on both delayed groups, "roughly proportional to the magnitude of the delay". However, they expressed that the time required to compensate for latency was "very slow". They extrapolated that it would take about 5h of exposure to latency for a complete compensation. Miall and Foulkes extended their work in a new experiment with a similar design but with five 1h sessions over 5 days. They showed that 5h exposure was still not sufficient to see complete compensation [21].

Results about the time needed to compensate for the latency seem contradictory. Rohde et al. argue that the unpredictability of the tracking task used by Foulkes and Miall may be a factor in latency compensation [29]. Rohde et al. experimented with a predictable vs. unpredictable indirect tracking task in a single session. They showed that only the group in the predictable condition shows compensation with a decrease of the tracking error with time. In a more recent work, Rohde and Ernst argue that adaptation might occur only when feedback delay is unequivocally identifiable "from other perturbations that can produce superficially similar effects, such as inertial changes, spatial offsets, etc..." [28].

Our experiment tests if the compensation, uncovered in the case of indirect interaction, also occurs in the case of direct touch. In indirect interaction, the user sees a unique feedback and the studied inconsistency is between the proprioceptive feedback of the hand and the visual feedback of the cursor which represents the hand. With direct touch, the user can see both the hand (which is not delayed) and the feedback (a dragged object for example) which suffers from the system latency. The inconsistency is thus between the expected behavior of the dragged object and its actual behavior. In this situation, learning to compensate for latency could happen differently.

Previous work also demonstrated that learning can take different durations depending on the nature of the task, we thus designed an experiment spread over several days in anticipation of a potential long-term compensation.

Transfer of learning

We also question if the compensation for latency developed for a given task is specific, or if it can be used on another task. In the motor control literature, this phenomenon is studied through the *transfer of learning* paradigm.

Learning a task leads to skill improvement and these modifications might have an impact when performing another task. When an improvement of performances is observed on the other task, a "transfer of learning" has occurred [27]. In this paper, we study if latency compensation, through training on a particular task, can transfer to another task with the same amount of latency.

De la Malla et al. studied the transfer of learning of delay compensation between different interception tasks [7]. They showed that the learning can transfer to tasks with little variations from the training task (like a starting position displacement) but does not seem to transfer across different interception tasks, even if they seem similar. For example, there was no evidence of transfer between a "target interception" task and a "passing through a moving gap" task. In this experimental design, participant's hand is hidden and users only see a cursor on the screen. Transfer could be different in the case of direct touch where the hand is visible.

This specificity of the transfer of learning was also observed in a recent HCI study and was used as a behavioral metric of the sensorimotor similarity between different interactions in sequential pointing [3].

EXPERIMENT: DEALING WITH TOUCH LATENCY BEFORE AND AFTER PRACTICE

Experiment design

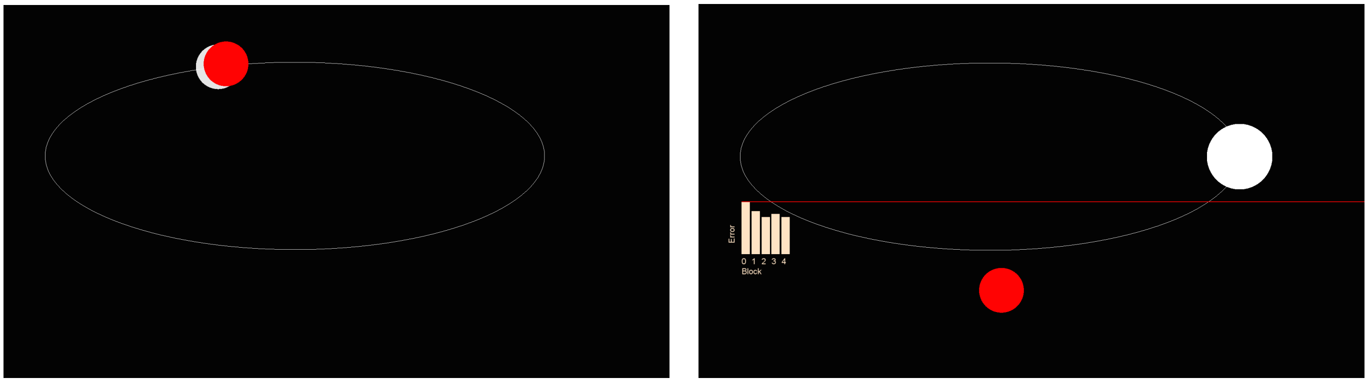
We design our experiment to test these two hypothesis:

- People can learn to deal with latency on a tracking task and, through practice, their performances become as good as when performing the task without latency.
- This learning can then be transferred to a similar task.

Participants perform a tracking task during ten practice sessions spread over two weeks. Our approach is to compare the effect of the latency on the task at first exposure and after the training. We expect to see an improvement of performances with practice. We also need to distinguish in this improvement which part accounts for the acquisition of skills specific to the task (learn the trajectory and speed of the target or how to track smoothly) and which part accounts for the compensation of latency. Thus, we used a *control group* trained on the same task but with minimal latency (i.e. 25 ms, as low as our system permits). The group dealing with latency is called the *test group*. During the practice sessions, the improvement of performances of the control group can only be due to the development of tracking skills. By difference with the test group, it enables to assess the improvement due to the learning of latency compensation. The test group, having to deal with latency, should show lower performances in the first session. Comparison between groups at the last session also enables to evaluate if latency can be totally compensated (performances of the test group finally catch up those of the control group), only partially or not at all.

The experiment imposes a between subject design: participants have to be tested in the same condition for two weeks, and their training then prevent their participation in another condition. The between subject design introduces a difficulty: the performance difference between the test and control group could be partly explained by a difference in the average participant performance of each group. We thus control the assignation of each participant to the test or control group by assessing their performances on a tracking task that they perform before the first session. The individual performance is computed on ten trials of tracking with the high level of latency used in the main sessions. We call this the *balancing task*. It is executed at the very beginning of the experiment. This procedure enables us to balance both groups in term of participant's abilities to track a target and to deal with latency.

The test group is trained with 75 ms of latency which approximates the latency of many current commercial devices [8]. In an ideal case, the control group should be trained with no or negligible latency. But a system offering touch interaction on a large surface with negligible latency has never been achieved. We approximate the zero latency condition by replicating the system of Cattan et al. [5]. This system uses a low latency



(a) Graphical output during a trial. The red disc is controlled by the participant's finger, the white disc (partially covered) is the target, which follows the ellipse path counterclockwise. Here, the target is not totally covered by the red disc and the crescent moon is in front indicating that the participant is late.

(b) At the end of a trial. The large white disc is the starting area for the next trial. A bar chart (displayed only between two trials) presents the results of the session's previous trials. The horizontal red line corresponding to the 0th trial is the averaged tracking error of the last session and sets a goal for participants.

Figure 1. The display at two different moments of a session

device coupled to a prediction technique (details are provided in the *Apparatus* section). We assume that the potential effect of the remaining latency in the control group is negligible compared to the effects of the 75 ms latency in the test group. For simplification purpose, in the remainder of this article we will refer to "with latency" for the test group, and to "without latency" for the control group.

To investigate if the compensation is task specific, the transfer of learning is assessed with a second tracking task, called the "transfer task". The transfer task follows a different path than the main task; it is performed on the first and on the last participant's session.

20 persons (5 females) participated in the study, all right-handed, with a mean age of 28 (range 23–37, sd = 3.74). Participants were mostly students coming from different labs/schools from our campus. Most of them own a smartphone and/or a tablet and use it every day.

Apparatus

Since we want to have a control group with participants performing the tracking task with a latency as small as possible, we need a touchscreen device that enables low latency. Literature introduces several examples of custom made devices with latency lower than current commercial devices [1, 5, 25]. The system used by Annett et al. and Ng et al. restricts the working area to a small surface of $24cm \times 16cm$ with a grayscale display only. We therefore choose to replicate the system used by Cattan et al. [5]. It implements touch detection on a regular screen using optical tracking. The system has a baseline touch latency of 25 ms that we regularly check using a predictive method [4]. Cattan et al. also used a linear prediction to virtually reduce the baseline latency of the system that we also replicate. This linear prediction is well suited for our tracking task since it implies no sudden change of speed. Cattan et al. extrapolate that with the prediction, the system is equivalent to a 9 ms or less latency system.

For the experiment, we use this 9 ms equivalent latency for the control group. For the test group, a 75 ms condition is simulated from the 25 ms baseline by adding 50 ms of idle time before dealing with received touch events.

Task

We choose a task difficulty that is high enough to allow some learning and performance improvement. We also choose a novel task so that all participants have no training on the task at the beginning of the experiment. In particular, we avoid the common "dragging an object to a target" task: most participants use touchscreens in their daily life, and thus they have various levels of training on this task with various levels of latency. A tracking task meets our requirements. In pilot tests, we observed that complex paths were not required to observe learning. Using a circle with constant speed, however, would not suit our objective. A constant speed translates to a constant spatial gap between the finger and the feedback. This is a special case which flaws the temporal nature of the latency and may, according to Rohde and Ernst [28], hide a learning effect. Therefore we choose an ellipse path with a target speed linked to the curvature by the 2/3 power law. According to Lacquaniti et al [19], this speed profile follows the natural motion of the hand.

Users control a red disc *object* (1.78cm radius) on the screen with the index finger. The screen displays the path followed by a white disc *target* (1.78cm radius). Displaying the path of the target ensures that the target motion is predictable which, according to Rohde et al. is a condition to have delay adaptation [29]. The target motion (position and speed) follows these equations:

$$\begin{cases} x(t) = 840 + 720 * cos(t) \\ y(t) = 640 + 270 * sin(t) \end{cases} \qquad \begin{cases} x'(t) = -720 * sin(t) \\ y'(t) = 270 * cos(t) \end{cases}$$

x and y are coordinates expressed in pixels for a screen size of 1920×1080 pixels, with a diagonal of 24 inches (pixel size=0.277mm), and with the origin at the bottom left. The time t is expressed in seconds. The target follows an ellipsoidal trajectory, making a complete cycle in 2π seconds. The ellipse is slightly off center to the top left to prevent that participants' right hand hits the bottom and right edges of the screen when performing the task.

Participants drag the object to a starting area at the rightmost point of the ellipse. The target area turns green to indicate that the object is inside. When the object is in the starting area for 0.5s, the target appears and begins its counterclockwise path on the ellipse. The participants' goal is to follow the target as close as possible. Since the target and the object have the same size, participant attempt to completely occlude the target with the object. A complete occlusion is difficult, and the target usually reveals itself as a white "crescent moon" under the red object, creating a clear visual feedback indicating if the participant is late (moon in front) or early (moon behind). The display during a trial is illustrated on Figure 1a.

During a *trial*, the target makes three laps. Error is not measured during the first lap to let the participant catch up with the target and get back the rhythm of the motion. During the two following laps, the tracking error is sampled at 120Hz: it is the distance between the disc center and the target center. A trial lasts around 20s ($3 \times 2\pi$). When the rightmost point of the ellipse is crossed at the end of the third lap, the trial ends. The trial error is recorded as the average tracking error over the two final laps. The system moves the object back to a fixed position at the bottom left of the screen, ready for the next trial.

We allow participants to follow their progress by displaying their trial error at the end of each trial, as illustrated on Figure 1b. The display is a bar chart where the first bar represents the participant's trial error averaged across the trials in his previous session. In the specific case of the first session, the first bar just indicates the first trial error. The first bar has always the same size on the different sessions and is emphasized by a horizontal red line. The following bars represent the error of the current session's trials. The size of each bar is computed relatively to the first one. Participants implicitly attempt to not exceed the red line: this means that they improved their performance compared to the previous session. The bar chart disappears as soon as the object is dragged to the starting area to begin the next trial.

We also define a second trajectory to study the transfer of learning. The study by De La Malla et al. indicated that transfer only occurs when tasks are very similar [7]. A task too dissimilar to the ellipse would thus certainly lead to a null transfer which would not be useful for interpretation. We prefer a very close task to see what amount of transfer can be observed in favorable transfer conditions. We choose the eight-shaped curve defined by the following equations:

$$\begin{cases} x(t) = 840 + 720 * cos(t) \\ y(t) = 640 + 405 * sin(2t)/2 \end{cases} \qquad \begin{cases} x'(t) = -720 * sin(t) \\ y'(t) = 405 * cos(2t) \end{cases}$$

The eight shape is interesting because the right side is very similar to the right side of the ellipse, the left side of the eight is similar to the ellipse but with the target going clockwise instead of counterclockwise. And finally, the center part is a diagonal trajectory in both directions that is not included in the ellipse. We thus can presume that the transfer of learning from the ellipse to the eight could be different on these different parts. The display during a trial for the transfer study is illustrated on Figure 2.

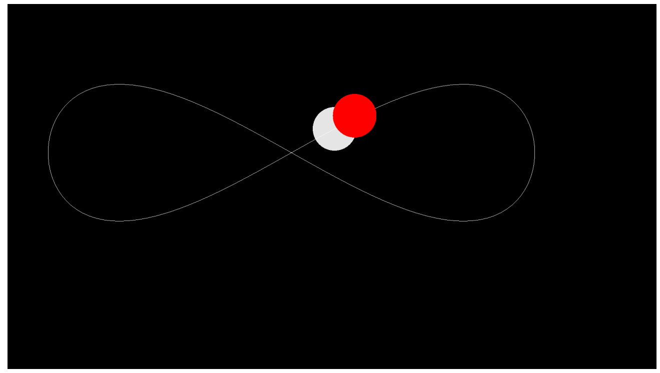


Figure 2. Graphical output during a trial of the transfer task.

Temporal distribution of training sessions

Sleeping has been shown to be an important factor in the improvement of motor skill performances [32]. We thus choose to spread our learning sessions on different days with the goal of observing clear learning effects. Gerbier et al. showed that an expanding schedule of repetition could optimize the retention [12] however we prefer to keep a rather flexible timetable for participants' convenience.

We choose to spread the complete training on the ellipsoidal tracking on ten sessions of sixty trials each. The session length is calibrated to last around half an hour. Pilot study showed that longer sessions may introduce a strong effect of fatigue. A semi-regular lapse of time between sessions is imposed: all sessions are done on different days (i.e. with at least a night of sleep between two sessions), two sessions cannot be separated by more than three days and the whole ten sessions have to fit within three weeks. In addition, participants perform their sessions around the same hour of the day. Eventually, due to schedule constraints, two participants of the test group did two of their ten sessions on the same day (separated by several hours). We do not observe any strong effect of the session distributions: the learning curves are similar between participants doing three sessions per week every other day and participants coming every day of the week (except week-ends).

During sessions, short pauses are *allowed* between each trial but longer 30s pauses are *imposed* every ten trials to prevent participants to rush through and be jaded for the last trials.

Our secondary objective is to study the potential transfer of learning between the ellipsoidal tracking and the eight-shape tracking. For this purpose, we measure the performance of participants on the eight-shape at the beginning of the first session and at the end of the last session. This is the *transfer task*. The distribution of sessions is summarized on Figure 3.

Measurement and analyzes

During the training part, 20 participants performed 10 sessions of 60 trials, totaling 12000 trials. From these, 34 trials (0.3 %)

Session	1			2	...	9	10	
Task	bal.	trans.	Training					trans.
Ctrl	10	10	60	60	...	60	60	10
Test	10	10	60	60	...	60	60	10

Figure 3. Summary of the temporal distribution of sessions depending on the group: control (Ctrl) or Test. "bal." is short for "balancing task" and "trans." for "transfer task". Numbers indicate the number of trials in each session. Grey cells are trials with latency, white cells are trials without latency. Sessions 2 to 9 are identical.

were removed as outliers due to optical tracking problems (the optical marker was somehow hidden during the trial).

We use the tracking error (i.e. the distance between the center of the dragged object and the center of the target) as the performance metric. We define *error*, the dependent variable of the experiment, as the average error per sample over all trials of a participant's session. *error* is used to assess improvement from training, but also for balancing the performances of the control and test groups, and for measuring the transfer of learning. In the case of the 10 trials, used to measure the transfer of learning in the first and last session, the 5 first trials are discarded because the abrupt switch from ellipse to eight shape results in very high *error*. We thus measure the transfer of learning after a quick (less than 2 min.) adaptation to the new shape.

We study the effect of two factors on *error*:

- SESSION is within-subject, with levels 1 to 10. In the specific case of the transfer of learning study, SESSION can only take two values, 1 and 10.
- GROUP is between-subject, with levels *control* and *test*.

RESULTS

Balancing task

Averaged *error* for the control group ($4.57mm \pm 0.63$) and the test group ($4.58mm \pm 0.70$) on the balancing task are very similar, indicating a good balance of user performance between the two groups. An unpaired t-test reveals a high probability of the NULL hypothesis ($t(18) = -0.04, p > .96$).

Training along sessions

error curves along sessions, averaged over groups, are presented in Figure 4.

There is a larger variability between participants in the test group than in the control group. It can be seen on Figure 4 with larger confidence intervals for the test group. Both groups showed approximately the same standard deviation between participants on the balancing task (0.63 vs. 0.70), hence this larger variability in the test group during the training demonstrates that people react differently when facing latency. Some of them have their performances badly impacted while others manage to compensate more quickly. On the contrary, in the control group, the participants followed more or less the same learning curve.

On the first session, performances are lower in the test group than in the control group, showing the negative influence of latency. There is a 18 % loss in performances due to latency

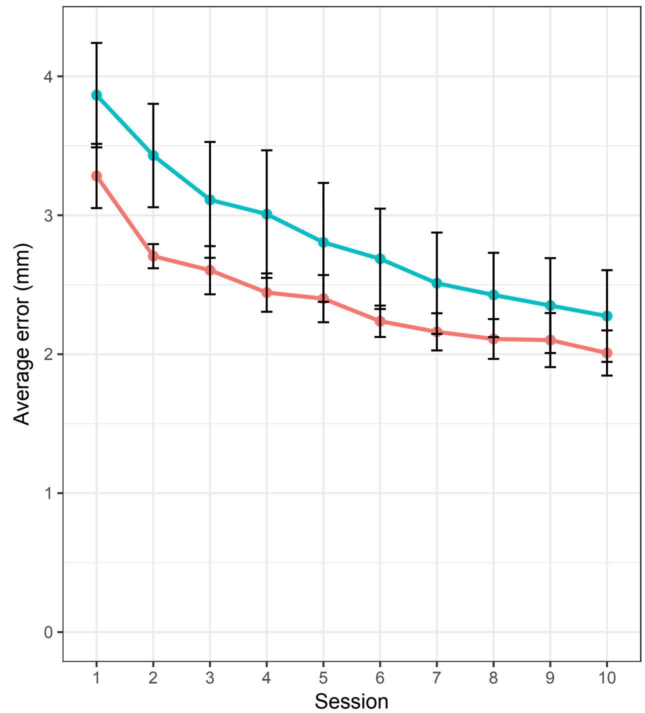


Figure 4. Tracking error in both control (red) and test (blue) groups averaged over the participants for each session.

(from 3.28mm to 3.86mm, a 0.58mm loss). While performance loss on target acquisition time has been well studied, this first result shows that latency also greatly impacts participants' tracking abilities.

On the ten practice sessions, a two-way ANOVA reveals a strong global effect of SESSION ($F_{9,162} = 125, p < .001$). We can see on Figure 4 that this effect is due to a continuous decrease of the tracking error with practice, showing the learning of new skills. All participants improve their performances after 10 training sessions: tracking error is significantly lower on the last session compared to the first session (paired t-test, $t(19) = 16, p < .001$). The intra session standard deviation (variation on the sixty trials of a single session) globally decreases from 0.51mm to 0.25mm showing that participants also improve their regularity through sessions. Participants do not have the same learning speed. Some of them present stabilized performances on the last three sessions whereas others are still making progress on the last session. Globally, 50 % of the decrease of *error* happens between the first and the third session. This is coherent with many previously observed learning curves, indicating that improvement becomes more difficult with time.

The ANOVA also uncovers an effect of GROUP on *error* ($F_{1,18} = 7.2, p = .015$). The test group, dealing with latency, has globally lower performances over the ten sessions. But the ANOVA reveals an interaction between GROUP and SESSION ($F_{9,162} = 3.7, p < .001$). We compare the performance difference between groups before and after the training. The 0.58mm performance loss due to latency, measured on the first day, decreases to 0.26mm on the last day. This is a 54 %

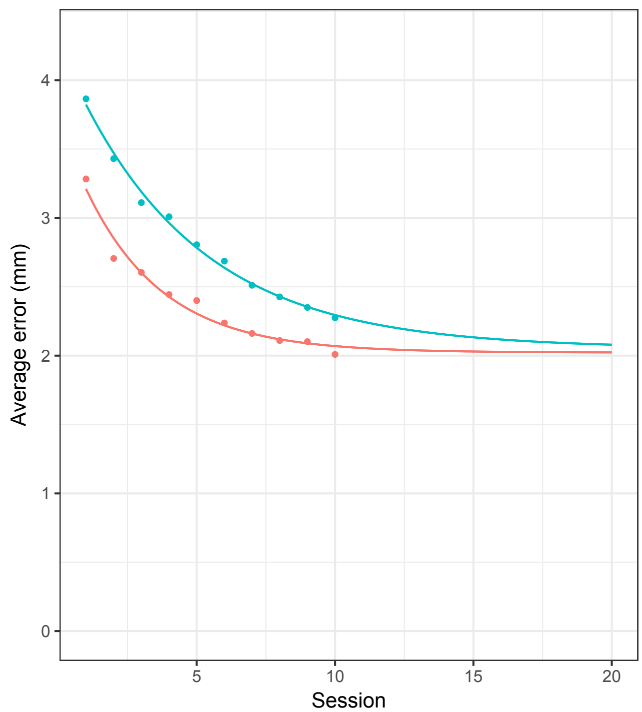


Figure 5. Extrapolation of the learning data for the control (red) and the test (blue) group with exponential models.

reduction. The negative influence of latency is thus more than halved in ten practice sessions. The test group is catching up progressively with the control group. As the only difference between groups is the presence of more latency in the test group, it shows that in addition to tracking skills, participants in the test group also develop specific mechanisms that enable them to compensate for latency.

We also run two post-hoc unpaired t-tests (with Bonferroni correction) on the first and last session which reveal a significant difference between the two groups in the first session ($t(18) = -2.98, p = .016$) but not in the last session ($t(18) = -1.64, p > 0.2$). These tests have to be interpreted cautiously though. The continuity of the curves shape in Figure 4 suggest that the difference in the last session is not random.

These results show that people can compensate at least partially for a delayed feedback on the tracking task. The strong negative influence of latency that is usually observed at initial exposure may not be as problematic as previously thought when considering long term exposure.

Curve extrapolation

The curves in Figure 4 did not reach a plateau. This suggests that more progress could be expected if participants had trained on more sessions. To get an idea of the final performances gap between the two groups, we experimented with curve extrapolations. Learning curves are known to be described by power or exponential laws [23]. After trying various models, we obtained a good fit with an exponential model.

Figure 5 shows the result of the exponential fit on the learning curves of both groups. The fit is done by a least square minimization of the error with the following equation (with three degrees of freedom): $y = Ae^{-Bx} + C$. The fit is very good in both conditions with $R^2 = 0.97$ for the control group and $R^2 = 0.99$ for the test group. With the models, the gap of performances between the groups continues to reduce after the tenth session until it reaches an asymptotic level. The convergence levels are 2.02mm and 2.05mm for the control and test groups, respectively. This extrapolation indicates that people dealing with latency are able, through practice, to almost entirely compensate for 75 ms of latency for a tracking task.

However, we tested fitting our model on a smaller number of sessions and we observed that the asymptotes of both groups systematically lower as we add further sessions in the fits. This indicates that despite the good fit, the exponential model may converge too quickly compared to the real learning curves. In other words, the model certainly still overestimates the actual asymptotes, even with 10 points for each group.

Transfer task

The results for the transfer task are illustrated on Figure 6.

We analyze *error* on the global eight-shaped trajectory in sessions 1 and 10 (Figure 6, "Global"). An ANOVA shows a strong effect of GROUP ($F_{1,18} = 25, p < .001$) meaning that the test group, dealing with latency, has a bigger *error*. It also shows a strong effect of SESSION ($F_{1,18} = 123, p < .001$) meaning that participants better perform the tracking task on the eight shape in the last session compared to the first.

In the control group, the averaged tracking error on the eight shape is reduced from 4.03mm on the first day to 2.88mm after the training on the ellipse. This 28 % reduction of the error is too large to be attributed only to the training offered by the 10 trials of the first session: the two sessions were separated by at least 10 days without training. Furthermore, the training between sessions 1 and 2 of the ellipse only leads to a 17.6 % improvement while offering six times more trials. Hence, this amplitude of improvement reveals a transfer of learning between the two tasks. The transfer appears as incomplete because the 28 % reduction is smaller than the 39 % error reduction that occurred on the ellipse (from 3.28mm to 2.01mm).

The ANOVA also uncovers an interaction between GROUP and SESSION ($F_{1,18} = 6.9, p = .017$) which indicates that the test group had more transfer than the control group. This indicates that not only the tracking skills are transferred, but also the skills which enable to compensate for the latency.

The error difference between groups, due to the latency, is 1.42mm on the first session (4.03mm and 5.45mm). It is larger than with the ellipse because the velocity profile on the eight shape is more variable and has higher peaks, which makes latency harder to tackle. The difference is reduced to 0.72mm (2.88mm and 3.60mm) in the last session. This is a 49 % reduction of the gap of performances between the two groups, similar to the 54 % found for the ellipse. This confirms that

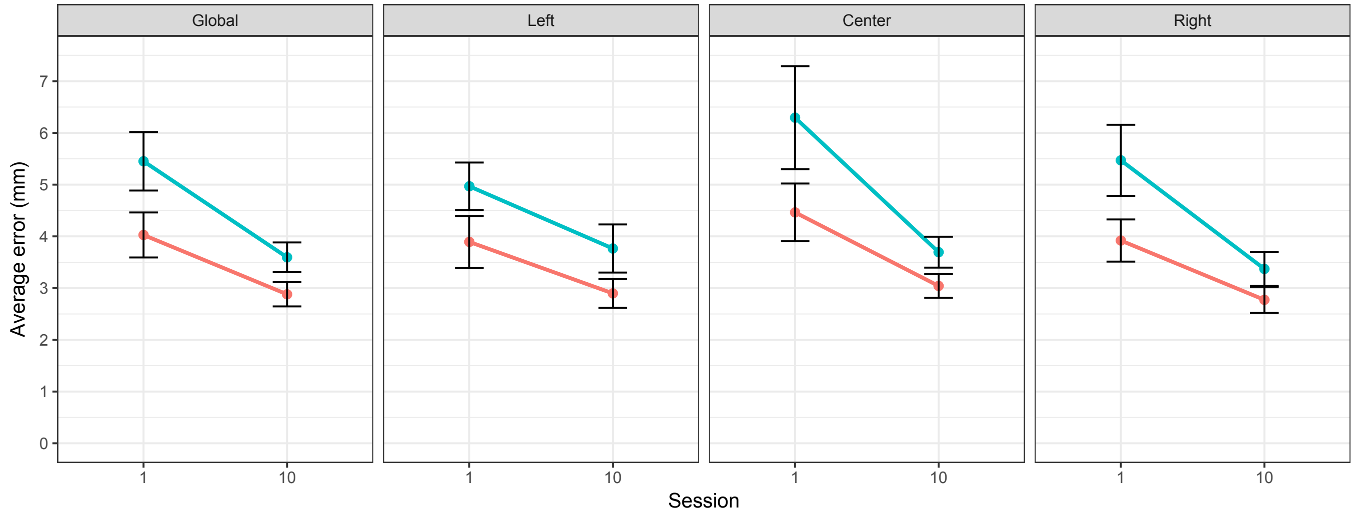


Figure 6. Transfer or learning. Tracking error in both control (red) and test (blue) groups averaged over the participants for the transfer task. The "Global" graph gives the transfer for the whole shape while the next graphs give the detailed transfer for different sections of the eight shape.

the ability to compensate for latency can transfer between the two tasks.

We tested our hypothesis that the transfer could be different on different sections of the eight shape. We cut the eight into three parts ($x < 600$, $600 < x < 1080$, $1080 < x$) and we included the factor EIGHT_PART with levels LEFT, CENTER and RIGHT. Figure 6, "Left", "Center" and "Right" shows the tracking error on the three parts for both groups in session 1 and 10.

In session 1, the tracking error is larger in the center part for both groups since in this area, the target is at its highest speed, making it harder to follow. In addition, participants of the test group had worse results on the right part compared to the left part: the position of the hand tends to hide the delayed feedback on the right part which makes the tracking more difficult.

The transfer of learning appeared to be strong for both groups in the center part, which is surprising since the trajectory is very different from the ellipse in this section. However, the amplitude of the tangential velocity is very similar. The ability to transfer tracking skills may thus be related to the velocity in a larger part than to the shape of the trajectory.

A triple interaction between GROUP, SESSION and EIGHT_PART indicates that the difference of transfer between the two groups is not the same for each eight part ($F_{2,36} = 3.5, p = .04$). The strong negative impact of latency in the test group on the center and right parts creates a lot of room for improvement, which may explain a bigger transfer in the test group than in the control group in these sections. On the center, the velocity of the target has also a bell-shaped profile, which is similar to the ellipse which certainly favors the transfer. On the contrary, on the left part, latency compensation does not seem to transfer: there is the same amount of transfer in both groups. This can be explained by the combination of two elements: a trajectory in an opposite direction compared to the ellipse and a velocity which cannot be found in the ellipse velocity profile.

These results indicate that the transfer of latency compensation for a tracking task can depend on both the trajectory shape but also the velocity profile of the target.

POST-EXPERIMENT: LEARNING RETENTION

Hypothesis and design

Our participants are able to greatly improve their performances after 10 training sessions on the course of 3 weeks. But would this persist over time? We ask to available participants of our study to come back between 7 and 9 weeks after their last training session. 12 participants (6 control - 6 test) are available for this post-experiment test. We check that, during this time, none of them had any exposure to tracking tasks similar to our experimental task. The task is equivalent to half a session of training (i.e. 30 trials). It is performed by participants in the same condition as in the training experiment.

On the control group, we expect participants' performances to drop a little compared to the tenth session due to the lack of practice. A similar drop is expected on the test group for the same reason. However, since latency is usually considered as a perturbation, our hypothesis is that the latency compensation would only be temporary. We expect participants of the test group to lose their ability to compensate for latency, and thus to observe a larger drop of performances than in the control group.

Since we need to compare the between-subject groups, we also verify that the two sub-groups of 6 participants are still balanced. The average errors of the two sub-groups are very close: 4.36 for the control group and 4.28 for the test group, the difference is not significant as revealed by an unpaired t-test ($t(10) = 0.3, p > .77$).

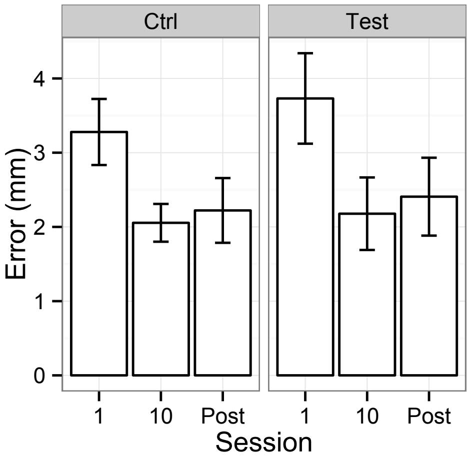


Figure 7. Error in sessions 1, 10 and in the Post-test for the Control (Ctrl) and Test groups.

Results

Since participants did not practice the task for at least seven weeks, the post session reveals a notable improvement during the session. We consider the first twenty trials to be a quick (less than 10min) re-familiarization with the task and we measure the tracking error on the ten last trials. Results are illustrated on Figure 7. An ANOVA is run to evaluate the effect of SESSION (10th vs. post-test) and GROUP on the error.

The tracking error on the post-session is significantly above the error level of the tenth session ($F_{1,10} = 6.39, p = .03$), but the loss of precision is minimal: +0.2mm. In particular, this is a small step back compared to the improvement achieved in the training ($-1.39mm$). There is no significant effect of GROUP ($F_{1,10} = 0.46, p > .5$) nor interaction between GROUP and SESSION ($F_{1,10} = 0.16, p > .69$). This means that participants are able, in less than 10 minutes, to reach a similar level of performances to that of their tenth training session. In addition, the drop of performance between the tenth and the post sessions is found similar in the two groups.

Our hypothesis of the temporary learning of latency compensation is thus invalidated: latency compensation allowed participant of the test group to acquire a sustainable amelioration of their baseline performance on the tracking task.

In the next section, we discuss the consequences of the results of the experiment and the post-experiment.

DISCUSSION

Latency compensation through practice

Considering only the control group, the results of the study show that, with practice, participants have learned skills related to tracking a target on a touchscreen. The tracking on the eight shape shows that these skills are not specific to the ellipse trajectory but are more general and can transfer to another trajectory with a different velocity profile.

The test group shows bigger improvements during the training, catching up progressively with the control group. This means that, in addition to skills related to tracking, participants of the test group also developed mechanisms specific to the presence of latency in order to compensate for it.

Different velocity profile on the ellipse and on the eight shape lead to different spatial gaps between the finger and the disc when dealing with latency. Relying only on visual feedback would give only spatial indications and thus would not lead to any transfer of latency compensation between the ellipse and the eight. The substantial transfer to the eight trajectory that we observed in the test group thus indicates that other mechanisms must have been developed by participants in the test group to integrate the temporal nature of the delay and to manage to apply them to another trajectory.

Latency seen as an inherent feature of touchscreens

Our experiment proved that users can learn to compensate for latency with practice. According to Krakauer, there are two different learning processes: "adaptation", and "skill learning" [18].

Adaptation is a modification of brain mechanisms in order to perform action accurately in the presence of a sensorial or mechanical perturbation. For example, it is well known that people adapt quickly their actions when facing a displacement of their field of view induced by prismatic glasses: with training, their performance in reaching or pointing tasks quickly attains pre-perturbation levels [30]. When the perturbation is removed, the observation of a negative after-effect shows that the user continues to behave as if the perturbation was still there. This after-effect is taken as an evidence of an adaptation: some visuomotor recalibration occurred in the users' brain.

"Skill learning" is defined as a shift in the speed-accuracy trade-off for a task when no systematic perturbation is present. In contrast to adaptation, which is the adjustment to a perturbation in order to get back to the pre-perturbation performances baseline, skill learning is the improvement of this baseline. Skill learning requires more time than adaptation. A more complete review of the learning processes can be found in Krakauer et al. [18].

Since touch interaction has been designed to mimic actions of the physical world, which does not suffer from latency, we assumed that a delay on a touchscreen device would be seen as a perturbation to which people could adapt. Participants indeed developed mechanisms to compensate for latency on our tracking task and their performances came close to those of a group trained with almost no latency. Adaptation, however, as defined in the motor control field, is certainly not the correct term to describe this progress. The learning retention experiment showed that the mechanisms developed to compensate for latency were still present after seven weeks without practice. This demonstrates that these mechanisms are more related to skill learning, i.e. the improvement of a baseline performance using a device with latency.

Considering the everyday use of touchscreens, when users deal with a laggy device, they may not see it as a perturbation, but rather as a tool that takes time to be accustomed with. In the motor control studies where the hand was not visible or the interaction indirect [6, 7, 10, 21], adaptation occurred. In the case of touch latency, the interaction is direct and it is as if the hand manipulates a (lagging) tool. Hence, observing skill learning is not surprising.

One way to consider latency is as an inherent feature of a touchscreen device. For an analogy, someone may learn to drive an old car with no power steering. This person may take more time to learn how to turn the steering wheel correctly compared to someone learning on a power steering car, but this will not be seen as a perturbation. And this person should be able to reach the same driving performance as the one who learned with power steering.

This view of latency is strengthened by the time required by the gap of performances to decrease. As stated by Krakauer, “Skill learning takes much longer than adaptation” [18]. In our study, the difference between the two groups did not disappear after one or two sessions but decreased progressively for the whole ten sessions. This may also explain some participants’ comments found in Ng et al. [25]: they report that users can be unaware of latency as long as the comparison is not explicit. But once participants have tried a very low latency device ($\approx$1 ms), they find the latency of current devices “completely unacceptable”. Before having experienced a better device, users may not have considered latency as a perturbation, but rather as an inherent feature of the touchscreen device.

Impact of touch latency on users performances

The development of new skills through 10 practice sessions resulted in the reduction of 54 % of the negative impact of latency on the tracking task. Furthermore, there is retention of these skills, even after more than seven weeks without training. Hence, our study indicates that the impact of touch latency on users’ performances is less severe than what could be expected from previous work.

We chose an experimental task with a *predictable* trajectory, and a task in which latency cannot be confused with other perturbations (spatial or inertial). The goal was to favor learning [29, 28]. More common HCI tasks will have to be studied to assess the generality of our results. Target acquisition tasks include motions that are strongly erratic during the adjustment phase [5], hence compensating for latency on such motion should be more difficult than for the smooth trajectory used in our work.

The transfer of learning between different levels of latency would be another interesting investigation since users interact with many devices and applications that exhibit different levels of latency. Such study could tell if the skills developed to compensate for the latency in one condition could propagate easily to different activities, or if new learning phases would be necessary. A lack of transfer between different levels of latency would argue for systems and applications with identical latency levels.

Finally, although this work indicates that the impact of latency may be less harmful than previously thought, this should not restrain the development of low latency devices: learning to compensate for latency takes a lot of time, and may only transfer to very similar tasks. More importantly, compensating for latency may have a cost. Users dealing with latency may be able to reach performances similar as with no latency but this might imply more concentration and cognitive load and induce fatigue. These effects should also be objectively evaluated in future studies.

CONCLUSION

We presented a study aimed at getting a better insight on the negative effects of touch latency on users’ performances. We demonstrated that users can learn new skills that enable them to compensate substantially for the negative influence of latency on a tracking task. The performance loss due to latency was more than halved with 10 half-hour training sessions. The retention of this skill was good more than seven weeks after the last training session. These results nuance the results of previous studies where the effect of latency was only tested on short sessions. Latency should still be considered as a major hindrance since learning to compensate for latency requires a sizable effort, it may not occur or transfer to every types of tasks, and it may have a high cognitive cost. All these will have to be studied in further experiments.

ACKNOWLEDGEMENT

This work has been partially supported by the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01).

REFERENCES

1. Michelle Annett, Albert Ng, Paul Dietz, Walter F. Bischof, and Anoop Gupta. 2014. How Low Should We Go?: Understanding the Perception of Latency While Inking. In *Proceedings of Graphics Interface (GI)*.

2. Gisa Aschersleben. 2002. Temporal control of movements in sensorimotor synchronization. *Brain and cognition* 48, 1 (2002), 66–79.

3. François Bérard and Amélie Rochet-Capellan. 2015. The Transfer of Learning as HCI Similarity: Towards an Objective Assessment of the Sensory-Motor Basis of Naturalness. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 1315–1324.

4. Elie Cattan, Amélie Rochet-Capellan, and François Bérard. 2015a. A Predictive Approach for an End-to-End Touch-Latency Measurement. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*. ACM, 215–218.

5. Elie Cattan, Amélie Rochet-Capellan, Pascal Perrier, and François Bérard. 2015b. Reducing Latency with a Continuous Prediction: Effects on Users’ Performance in Direct-Touch Target Acquisitions. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*. ACM, 205–214.

6. Douglas W Cunningham, Vincent A Billock, and Brian H Tsou. 2001. Sensorimotor adaptation to violations of temporal contiguity. *Psychological Science* 12, 6 (2001), 532–535.

7. Cristina de la Malla, Joan López-Moliner, and Eli Brenner. 2014. Dealing with delays does not transfer across sensorimotor tasks. *Journal of vision* 14, 12 (2014), 8.

8. Jonathan Deber, Bruno Araujo, Ricardo Jota, Clifton Forlines, Darren Leigh, Steven Sanders, and Daniel Wigdor. 2016. Hammer Time!: A Low-Cost, High Precision, High Accuracy Tool to Measure the Latency of Touchscreen Devices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2857–2868.

9. Jonathan Deber, Ricardo Jota, Clifton Forlines, and Daniel Wigdor. 2015. How Much Faster is Fast Enough?: User Perception of Latency & Latency Improvements in Direct and Indirect Touch. In *Proceedings of ACM CHI*.

10. Alexander J McC Foulkes and R Chris Miall. 2000. Adaptation to visual feedback delays in a human manual tracking task. *Experimental Brain Research* 131, 1 (2000), 101–110.

11. Sebastian Friston, Per Karlstrom, and Anthony Steed. 2014. The Effects of Low Latency on Pointing and Steering Tasks. (2014).

12. Emilie Gerbier, Thomas C Toppino, and Olivier Koenig. 2015. Optimising retention through multiple study opportunities over days: The benefit of an expanding schedule of repetitions. *Memory* 23, 6 (2015), 943–954.

13. Niels Henze, Markus Funk, and Alireza Sahami Shirazi. 2016. Software-reduced touchscreen latency. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, 434–441.

14. Zenja Ivkovic, Ian Stavness, Carl Gutwin, and Steven Sutcliffe. 2015. Quantifying and mitigating the negative effects of local latencies on aiming in 3d shooter games. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 135–144.

15. Ricardo Jota, Albert Ng Ng, Paul Dietz, and Daniel Wigdor. 2013. How Fast is Fast Enough? A Study of the Effects of Latency in Direct-Touch Pointing Tasks. In *ACM CHI*. ACM.

16. Topi Kaaresoja, Stephen Brewster, and Vuokko Lantz. 2014. Towards the temporally perfect virtual button: touch-feedback simultaneity and perceived quality in mobile touchscreen press interactions. *ACM Transactions on Applied Perception (TAP)* 11, 2 (2014), 9.

17. Mirjam Keetels and Jean Vroomen. 2008. Temporal recalibration to tactile–visual asynchronous stimuli. *Neuroscience letters* 430, 2 (2008), 130–134.

18. John W Krakauer and Pietro Mazzoni. 2011. Human sensorimotor learning: adaptation, skill, and beyond. *Current opinion in neurobiology* 21, 4 (2011), 636–644.

19. Francesco Lacquaniti, Carlo Terzuolo, and Paolo Viviani. 1983. The law relating the kinematic and figural aspects of drawing movements. *Acta psychologica* 54, 1 (1983), 115–130.

20. I Scott MacKenzie and Colin Ware. 1993. Lag as a determinant of human performance in interactive systems. In *Proceedings of ACM CHI*. 488–493.

21. RC Miall and JK Jackson. 2006. Adaptation to visual feedback delays in manual tracking: evidence against the Smith Predictor model of human visually guided action. *Experimental Brain Research* 172, 1 (2006), 77–84.

22. Mathieu Nancel, Daniel Vogel, Bruno De Araujo, Ricardo Jota, and Géry Casiez. 2016. Next-Point Prediction Metrics for Perceived Spatial Errors. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, 271–285.

23. Allen Newell and Paul S Rosenbloom. 1981. Mechanisms of skill acquisition and the law of practice. *Cognitive skills and their acquisition* 1 (1981), 1–55.

24. Albert Ng, Michelle Annett, Paul Dietz, Anoop Gupta, and Walter F. Bischof. 2014. In the Blink of an Eye: Investigating Latency Perception During Stylus Interaction. In *ACM CHI*. ACM, 1103–1112.

25. Albert Ng, Julian Lepinski, Daniel Wigdor, Steven Sanders, and Paul Dietz. 2012. Designing for low-latency direct-touch input. In *Proceedings of ACM UIST*. ACM, 453–464.

26. Andriy Pavlovych and Carl Gutwin. 2012. Assessing target acquisition and tracking performance for complex moving targets in the presence of latency and jitter. In *Proceedings of the 2012 Graphics Interace Conference*. Canadian Information Processing Society, 109–116.

27. David N Perkins, Gavriel Salomon, and others. 1992. Transfer of learning. *International encyclopedia of education* 2 (1992), 6452–6457.

28. Marieke Rohde and Marc O Ernst. 2016. Time, agency, and sensory feedback delays during action. *Current Opinion in Behavioral Sciences* 8 (2016), 193–199.

29. Marieke Rohde, Loes CJ van Dam, and Marc O Ernst. 2014. Predictability is necessary for closed-loop visual feedback delay adaptation. *Journal of vision* 14, 3 (2014), 4.

30. Reza Shadmehr, Maurice A Smith, and John W Krakauer. 2010. Error correction, sensory prediction, and adaptation in motor control. *Annual review of neuroscience* 33 (2010), 89–108.

31. Chess Stetson, Xu Cui, P Read Montague, and David M Eagleman. 2006. Motor-sensory recalibration leads to an illusory reversal of action and sensation. *Neuron* 51, 5 (2006), 651–659.

32. Matthew P Walker, Tiffany Brakefield, Alexandra Morgan, J Allan Hobson, and Robert Stickgold. 2002. Practice with sleep makes perfect: sleep-dependent motor skill learning. *Neuron* 35, 1 (2002), 205–211.