# Finding the right way for interrupting people improving their sitting posture

Michael Haller[1,3], Christoph Richter[1], Peter Brandl[1], Sabine Gross[1], Gerold Schossleitner[2], Andreas Schrempf[2], Hideaki Nii[3], Maki Sugimoto[3], Masahiko Inami[3]

[1] Media Interaction Lab, Upper Austria University of Applied Sciences, Austria
[2] Medical Technology, Upper Austria University of Applied Sciences, Austria
[3] Keio-NUS Cute Center, Singapore/Japan
`mi-lab@fh-hagenberg.at`

**Abstract.** In this paper, we present three different ways of interrupting people to posture guidance. We developed an ergonomically adjustable office chair equipped with four sensors measuring the office worker's posture. It is important that users do some training after bad posture and be alerted of this; therefore, we implemented three different alert modalities (Graphical Feedback, Physical Feedback, and Vibrotactile Feedback), with the goal to find out which of the techniques is the most effective interruption modality without causing a huge disruption effect. To measure the task-performance, we conducted a formal user study. Our user study results show there are different effects on performance and disruptiveness caused by the three interruption techniques. While the vibrotactile feedback might have higher information awareness benefits at the beginning, it causes a huge intrusion side-effect. Thus, the physical feedback was rated less disruptive to the workflow as the other two feedback modalities.

**Keywords:** Posture Care, Interrupts, Physical Feedback, Graphical Feedback, Vibrotactile Feedback

## 1    Introduction

The typical office workplace is a computer workstation where the average office worker spends about 50,000 hours seated in the course of his/her working life. As a consequence of this motion deficiency about 40% of all office workers are going to have back problems [10]. The computer work typically results in low-level static loading of back, shoulder and neck muscles as well as in static spine loading. Prolonged, static sitting is associated with muscular disorders, the development of disc degeneration or the increase of spine stiffness [6,15]. Dynamic sitting, thus a behavior where the person dynamically activates lumbar muscles and spine movement is increased in a physiologically reasonable range is beneficial and will help to prevent back pain.

   For most people it is difficult to check their sitting behavior and posture while working concentrated on a task and most of them usually are not aware of the time they spend on work activities while sitting *statically* [9]. In order to improve the sitting behavior of office workers, we developed an intelligent office chair, which

allows us to classify the sitting posture and the corresponding time the person sits statically in any position (**Fig. 1**).



**Fig. 1.** Sitting postures that can be classified by the intelligent office chair. 1. upright, 2. leaning back, 3. leaning forward, 4. sitting at the front edge, 5. leaning right, 6. right leg crossed over left leg, 7. left leg crossed over right leg, 8. slouching.

The aim of the intelligent office chair is to guide the person through an effective feedback to a more dynamic and healthy sitting behavior. Our setup is based on a regular adjustable office chair, equipped with four independent, especially designed force transducers. The four sensors are located at each corner under the seating surface, thus making it possible to compute the coordinates of Center of Pressure (CoP). In order to do this, the reference frame is located in the center of the seating surface. The coordinates of the *CoP* vary according to the posture the person is sitting on the chair, which in turn allows users to classify the sitting posture (**Fig. 1**) and the time spent in the corresponding position.

After inadequate sitting, our system provides feedback [9] and the office worker gets an alert for performing a training session. We implemented three different techniques, including a graphical, a physical, and a vibrotactile interruption feedback, with the goal to find out which of the techniques is the most effective interruption modality without causing a huge disruption effect (cf. **Fig. 2**). Humans have their cognitive limitations, which makes them susceptible to errors once interrupted [2]. On the other side, we know from multiple resource theory that humans are able to perform different tasks in parallel as long as the tasks do not utilize the same cognitive resource [21]. Based on this theory, we assumed that office workers (who rely heavily on visual processing) would find both the graphical and physical feedback alert more distracting and less acceptable than the vibrotactile feedback. On the other side, the vibration might be harder to detect at the beginning, but it might be also harder to ignore once it was present [2]. We expected that the vibrotactile feedback followed by the physical avatar feedback is less disruptive than the digital

interruption modality. However, the digital technique might be the fastest way to get the users' attention, because it is directly shown in the user's field of view.
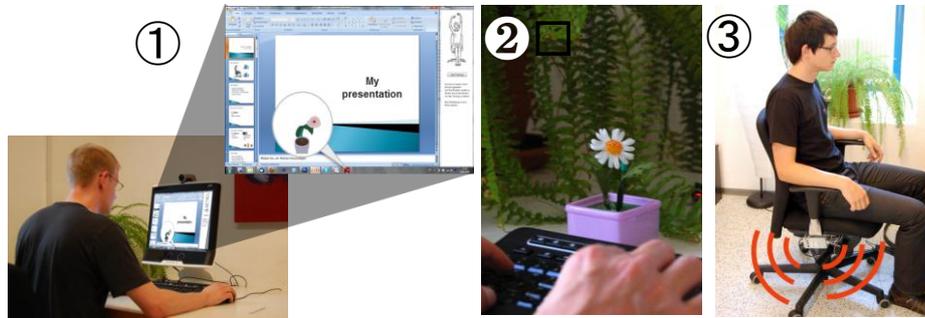


**Fig. 2.** Graphical (1), Physical (2), and Vibrotactile Feedback (3) should alert users to perform a training session.

Summarizing, all three techniques are being used with *two* different meanings: *providing a feedback* about the office worker's posture *and* triggering an *alert message* once the user should perform a relaxing training session. All techniques allow users to decide themselves when they want to switch from the primary task to the training session without capturing the user's context and realizing that they are not in a critical phase of a task.

## 2    Related Work

### 2.1    Posture Detection and Posture Guidance

There are three possibilities to train users sitting ergonomically correct on chairs, including a) an ergonomic chair with instable support, b) triggering a training session, and c) providing direct guidance on the actual sitting position. First of all, an ergonomic chair with instable support can be used, where users always have to balance their body on the chair that keeps them active (e.g. Haider - Bioswing) and/or by triggering users to perform a training session (hereby users can use a regular office-chair[1]).

WorkPace[2] is one of the most well known applications for users to train their muscle fatigue and recovery. The application alerts users whenever exercises (e.g. stretching exercises) should be performed. In contrast to WorkPace, we are tracking the chair during a longer period, thus getting permanent feedback from the user. Consequently our software can trigger alerts more precisely and provide optimized training exercises.

Another program is RSI-Shield[3], a user customizable application, which simply generates break events with a pre-defined frequency. During the break the user is

---

[1] http://www.rsi.org.uk

[2] http://www.workpace.com

[3] http://www.rsishield.com

advised to perform simple exercises, which can be done using a normal office chair. In contrast to the latter applications, the aim of our work is to detect whether the user is sitting correct or not and only to interrupt the user, if an unhealthy sitting position is recognized by the sensors of the intelligent chair.

Finally, Zheng and Morell [22] propose an ergonomic chair, which guides office workers to sit in a pre-defined position by providing a vibrotactile feedback. Force sensors placed on the sitting support as well as on the back rest of the chair are used to compute a feedback signal by using small vibration actuators. If the user is sitting in a desired position, one or more actuators vibrate in order to direct the user away from an undesired position. In their paper, the authors postulate that the sporadic "buzzes" helped users successfully to guide them into the desired posture. In contrast, our approach is to detect the sitting position of the user and the corresponding time the user sits in this position. Only if the user sits statically longer than a pre-defined time period, an interrupt will be generated. Since e.g. a slouching position is more harmful to the spine than leaning back using the backrest (compare position 8 vs. position 2 in **Fig. 1**) the according time period has been adjusted accordingly.

## 2.2    Interruption & Feedback

In the last decade, a number of research groups have presented a lot of work around interruption and recovery with the goal of having a highly efficient interrupt with low intrusion [4, 7, 8, 12, 13, 14, 20].

Different feedback modalities for efficient interruptions have been explored by Selker and Arroyo. In [3], the authors present five different feedback modalities including smell, heat, sound, vibration, and light. In their paper, they come to the conclusion that using the right channel can evoke certain memories, which again might be optimal to be used in a system.

The implementation of the Graphical Feedback has mainly been influenced by Bailey et al. [5], where the authors propose a new way interrupt people providing an optimal balance of information awareness with very low intrusion. However, they embed the alert window in the browser, which limits its usage for everyday applications.

In our setup, we also used a vibrotactile feedback to posture guidance. This feedback modality was mainly inspired by Zheng and Morell [22]. In their paper, the authors postulate that the sporadic "buzzes" helped users successfully to guide them into the desired posture. However, the authors did not compare the haptic feedback with other modalities. From the multiple resource theory we know that non-competing channels might have less negative disruption effects to office workers. In this paper, we went one step ahead and compared the impact of haptic and graphical feedback on the workflow.

## 3    Feedback Modalities

### 3.1    Graphical Feedback

The Graphical Feedback is based on the ideas of Bailey et al., where the authors postulate an "Adjusting Window" technique for notifying users of new information [5]. Using this technique, the primary (task) window shrinks and the alert window is getting docked to the right side of the main window.



**Fig. 3.** Additional alert messages can be visualized by shrinking the main window (*left*) and by docking a new window on the right side of the main window (*center*). Users can decide the right timing for performing the training (*right*). In this case the main window shrinks again and becomes the secondary window docked to the main window of the training window.

Bailey et al. postulated in [5] that the Adjusting Window technique seems to be an optimal compromise notifying users about new information with very low disruption. In addition, a changeable icon in the taskbar represents the current status of the users' posture. A small bending plant visualizes the posture of the user. In our setup, the main working window slightly shrinks, and the new alert window is wrapped along the right side of the primary window, see **Fig. 3** (*left, center*). Based on the size of the actual working window, the docked window is getting its size. The main window stays remained in the adjusted state until the user is performing the training session, thus clicking to "performing exercise button" inside of the docked window. Once done by the user, the exercise window is becoming the new main window and the old main window animates to a small window, docked along the left side of the exercising window, cf. **Fig. 3** (*right*). A one-minute exercise video is presented to the office workers, which they should follow. Finally, once the exercise has been performed, the main window is animated back to its original size and the small docked window on the right side disappears again. In contrast to Bailey et al.'s approach, we did not embed the docked window inside one specific application (e.g. browser), but we resized the actual working window, by getting the active window handle from the top-used window [16].

### 3.2    Physical Feedback

In contrast to the Adjusting Window technique, which is working only digitally, we also implemented a physical avatar (see **Fig. 4**).
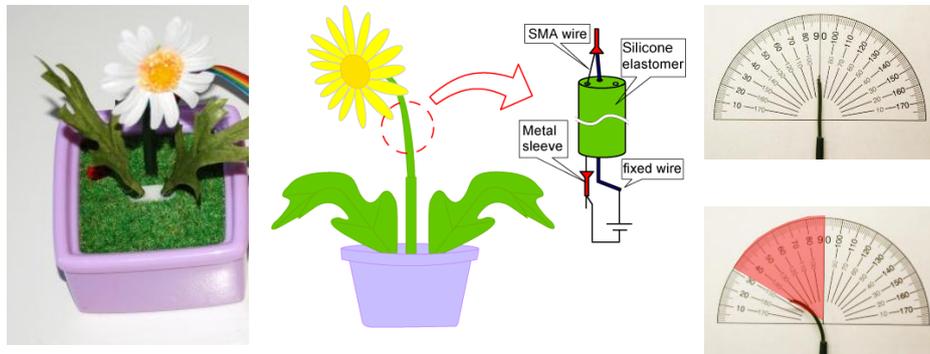
**Fig. 4.** *(left)* The bending plant represents the office worker's posture. (*center*) The image of the plant and detail of the actuator which consists of a fixed and a SMA wire. (*right*) The bending shape of the actuator without the bloom (*top:* without power, *bottom:* with power).

Instead of using a physical puppet [9], we used a plastic plant based on the Hanappa toy[4]. The original Hanappa plant, manufactured by Sega Toys, flexes its petal and/or leafs based on human's speech input. We modified the plant in three different ways:

- **Connection over USB:** while the original plant is a stand-alone plant without any bi-directional communication, we embedded a computer board to drive the actuator via USB.
- **Replacement of the actuator module:** The physical avatar uses a Shape Memory Alloy (SMA) technology, which makes it possible to bend its shape. Similar to the changing icon in the Windows taskbar, the plant can bend its shape to represent the user's posture. Moreover, it can shake itself to motivate the user to perform a training session. In the modified version, we replaced the actuator module to improve both the bending angle and time. The original Hanappa can bend its leafs and petal with an angle ±10°. We changed the SMA by using a longer Biometal wire with a larger diameter (Biometal BMF-100 for the petal and Biometal BMF-75 for the leafs). Both Biometal wires are able to change their length for 5% (thus the wire gets larger and/or smaller) depending on the applied power. Therefore, the modified version is able to bend with an angle of 60°, so that users get the impression that the plant is flabby and droopy. Similar to the changing icon in the Windows taskbar, the plant can bend its shape representing the users' posture. Moreover, it can shake itself once the user should get motivated to perform an exercise.
- **Adding another leaf:** Finally, we added another (second) SMA-leaf to the original Hanappa, which can be controlled again via USB. While the bloom of the flower should represent the user's head, both leafs are representing the arms.

---

[4] http://www.segatoys.co.jp/hanappa

### 3.3 Vibrotactile Feedback

Finally, we developed vibrotactile feedback with the aid of an actuator, so that users are motivated to change their seating behavior. To provide feedback about the status of the users' posture, they receive innocuous vibrations, along with sporadic "buzzes". The vibrations are created by the force feedback unit of a Logitech Rumblepad 2. In order to alert the user about a wrong sitting position, we used 0.5 seconds lasting „buzzes". The alert's magnitude was increased whenever participants constantly ignored the feedback. We started with a light vibration using 30% of the feedback's maximum strength (which was defined by the maximum force that the Rumblepad 2 could achieve) and increased the force with 6 discrete steps (40%, 50%, 60%, 70%) to finally 80%.

## 4 User study

### 4.1 Pilot-Study

In a pilot-study with 6 participants we tracked the participants' posture without any alert during a document-editing task. All participants had to extract words out of a text, where all spaces were removed. During the one-hour tracking session episodes of static sitting were identified, where the coordinates of the *COP* remained within a pre-defined region. During the one-hour session 5.8 ($SD$ = 3.9) episodes of static sitting were identified lasting for 7.2 ($SD$ = 13.5) minutes. The condition for an interrupt alert was satisfied, if an episode of static sitting lasts longer than 5 minutes. We found out that participants were achieving 5.8 ($SD$ = 4.9, $MIN$ = 2, $MAX$ = 13) of "silent" interrupt alerts during the one-hour tracking session.

### 4.2 Experimental Design

12 participants from the local software park were recruited to perform the laboratory user study. Participants were asked to perform three time consuming tasks. The experiment consisted of three different tasks, including the editing of a document, writing a transcript based on a video clip, and searching & planning a trip task. Summarizing, the study was a 3 (task) x 3 (feedback), counterbalanced within-subject design, which took 1.5 hours (10 minutes for each task). In the study, we measured the participant's posture for triggering an alert. In addition, the results of the pre-study motivated us to interrupt participants within 10 minutes at least *once* - even if they wouldn't have triggered an alert themselves (which actually did not happen). After each interrupt, participants could decide when to start a training session. If they ignored the training request for more than 15 seconds, the alert stopped and reminded them again after 30 seconds. This reminding sequence was repeated until the user finally started the training.

### 4.3 Tasks

Building on the experimental task classes suggested in [1], we devised three types of tasks for this study. The first type was a document-editing task. Articles from Wikipedia were converted into Microsoft Word document and shortened to an average of 1,820 words. Afterwards, spelling errors were introduced, words were replaced or skipped and some punctuation was removed. In all cases mistakes were indicated by the inbuilt spellchecker and marked with comments. Three instances of the editing task were created, building on similar yet distinct articles. Participants were instructed to make the requested changes as quickly and accurately as possible. This task required work within a single application.

The second set of tasks consisted of three news media clips. The video clips were about 2 minutes 20 seconds in length. Participants were instructed to produce a transcript of the narrator's text in Microsoft Word; thus participants had to work with two different desktop applications in parallel.

The third type of task was a combined web-search and planning task. Participants were asked to plan for a short trip and search for information regarding transportation and accommodation on the web in line with the key data given in the assignment. Participants were asked to document their findings in a Word document. Destinations, types of accommodation and means of transportation were varied between task instances to counteract strategic learning effects. Thus, in the third task, participants had to work with multiple applications and make active use of information from various web sources.

These tasks were chosen to cover a variety of task demands, while still being meaningful examples of commonly performed everyday tasks. The tasks are supposed to induce a variation of cognitive loads on different mental resources, triggering different cognitive processes and differing in complexity. To ensure that time on task was the same for all subjects, the length of the texts and videos as well as a the key data for the short trip were adjusted so that full completion of the task without interruption would be quite unlikely within the time limits set.

### 4.4 Participants

Twelve participants (3 females) aged between 22 and 42 years of age (average age was 28 years old) joined the user study. All participants had a good experience with both Microsoft Windows and their Office products.
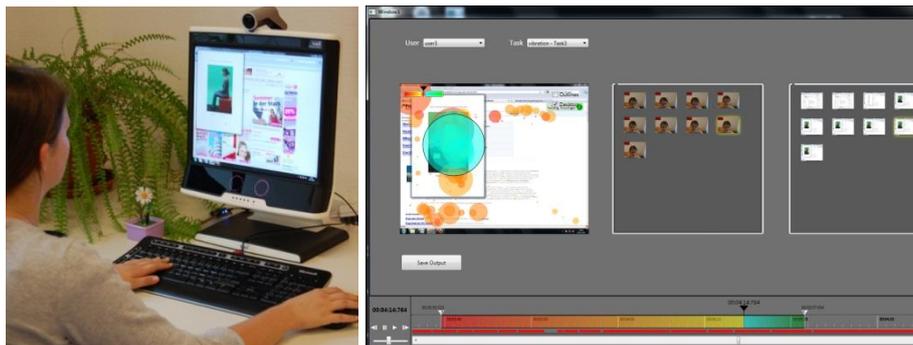
### 4.5 Apparatus

**Fig. 5** *(left)* depicts the apparatus of our study, including a Tobii eye-tracking screen, a webcam on the top for tracking the participants' faces and the physical Hanappa, placed on the right side of the LCD. All three feedback modalities (as described before) were used in the study based on the same hardware. The experiments were performed using a 17" TFT with a screen resolution of 1280×1024 pixels.

### 4.6    Procedure

Participants were welcomed and introduced to the purpose of the study. They were then given instructions on the tasks they have to perform. In addition, they were also informed that they would be interrupted periodically to do a training session. The participants were told to do the task exercises as fast and as accurate as possible. After each task they were shown a modified NASA-TLX survey.

### 4.7    Performance Measurements

We counted the number of training sessions that were postponed during a task. A training session was considered to be postponed if the participant did not react to the alert within 15 seconds. After every 30 seconds, the participant was reminded again to perform the training. If ignored, the exercise was again counted as postponed. Furthermore, we measured the times until an exercise was started after the initial alert. If an exercise was postponed, we added the time to the overall time until the exercise has been started. Finally, we measured the time for returning to the suspended primary task after a performed exercise. Therefore, we logged the first mouse and/or keyboard input event on the primary window after the exercise was finished. To confirm the measured times, we additionally double-checked the transition times from the end of a training to the resumed primary task through the Tobii's eye capturing screen[5]. In addition to the gaze data, the system logged any input on the screen, thus allowing us to exactly analyze the time when a user started to work on the primary task again. For further exploring the participants' behavior to an interrupt, we implemented our custom analyzing tool (cf. **Fig. 5** (*right*)). The tool parses Tobii's log files and allows visualizing gaze data over time, choosing a user-specific time interval. The timeline provides a time span control that helps to analyze the time period shortly before and after an interrupt happened. In addition to the recorded gaze data, the tool allows to browse through snapshots of the participants that were taken by the additional webcam, mounted on top of the Tobii LCD.



---

[5] http://www.tobii.com

**Fig. 5.** *(left)* The apparatus of the user study including the eye-tracking LCD of Tobii. *(right)* The custom analyzing tool allows a better analysis of how participants react to an interrupt. Capturing the participants' screen augmented with the gazing plot helped us to analyze better the results.

In order to provide additional information, we color-coded the gazing plots for periods before and after the interrupt. **Fig. 5** *(right)* depicts the timeline with an interrupt occurring after 29 seconds. The gazing blobs of the left screen of the figure occurred *before* the interrupt are visualized with warm gradient colors (red to yellow). In contrast, all gazing blobs *after* the alert are cold color-coded (turquoise to green).

### 4.8     Emotional State Measurements

To measure the effect of the interruptions on the user experience and emotional state, a modified version of the NASA-TLX survey way administered to the participants after each trial. While the NASA-TLX survey [11] was originally meant to assess the subjective workload, its scales are also relevant to the experience of interruption [1]. A particular advantage of the NASA-TLX is its short length with 6 items in the original and 8 items in our modified version, which allows us to present it frequently as required in this study.

The modified version used in this study was derived from the German translation of the TLX by [18]. While the Physical Demand scale was skipped, we added the following three items to get more specific information on the perceived impact of the continuous feedback and interruptions:

1. **Workflow:** How disturbing was the alert for the workflow?
2. **Feedback:** How disturbing was the continuous feedback?
3. **Training:** How disturbing was the alert to perform the training?

As suggested by [1], we administered the survey on paper rather than in electronic form to avoid interference with the experimental tasks.

## 5     Results

### 5.1     Performance Measurement Results

The results of the performance measurement are depicted in Table 1 (top/bottom). A two-way within-subjects analysis of variance (ANOVA) was conducted to evaluate the effect of feedback conditions and task type on the number of postponements, the time span from the training alert to the start of the training as well as the time it took to resume the main task. These results are summarized in Table 1 (top). For all tests an alpha level of 0.05 was used. There were no significant interaction effects. The Greenhouse-Geisser correction was used when the assumption of sphericity was violated. Post-hoc analyses were conducted on the significant main effects. These consisted of paired-samples $t$-tests with familywise error rate controlled across the test

using Holm's sequential Bonferroni approach. Significant differences between the means of pairs of conditions are presented in Table 1 (bottom).

| Performance Measure | Effect | F | p |
|---|---|---|---|
| Number of postponements | Task | $F_{2,22} = 3.404$ | 0.052 |
| | Feedback | $F_{2,22} = 2.566$ | 0.340 |
| Time until 1st training was started | Task | $F_{2,18} = 2.260$ | 0.133 |
| | Feedback | $F_{2,18} = 3.489$ | 0.052 |
| Time to resume to main task | Task | $F_{2,18} = 2.753$ | 0.091 |
| | Feedback | $F_{2,18} = 5.477$ | 0.034 |

| Performance Measure | Pair | F | p |
|---|---|---|---|
| Number of postponements | Edit – Transcribe | $t(11) = -2.359$ | 0.038* |
| | Edit – Search & Plan | $t(11) = -2.213$ | 0.049* |
| Time until 1st training was started | Graphical – Vibrotactile | $t(10) = 3.062$ | 0.012 |
| Time to resume to main task | Graphical – Physical | $t(9) = -3.443$ | 0.038* |
| | Graphical – Vibrotactile | $t(10) = -4.748$ | 0.003 |

Table 1. (*top*) Main effects for performance measures. (*bottom*) Significant mean differences along performance measures between pairs of conditions. Starred results indicate marginally significant results $0.05 > p > 0.0167$.

### 5.1.1 Type of Task

The type of task had a significant main effect on the tendency to postpone a training session. The fact that less trainings were postponed during the editing task ($M = 1.63$, $SD = 1.76$), in comparison to the transcription task ($M = 2.67$, $SD = 2.12$) and the search and planning task ($M = 2.79$, $SD = 1.68$), might be explained by the fact that this task required participants to work on a single document only, thus making it easier to respond to the interrupt and return to the main task afterwards. For the other two tasks, multiple documents had to be handled simultaneously which made a switch to the training and return to the main task more complex.

### 5.1.2 Type of Feedback

Even though the main effect for the type of feedback was non-significant, the tendency to postpone a training session appeared to be lowest under the vibrotactile feedback condition, cf. **Fig. 6** *(left)*. This might be due to the fact that the vibrotactile feedback was assessed as the most disturbing one (see results of NASA-TLX).
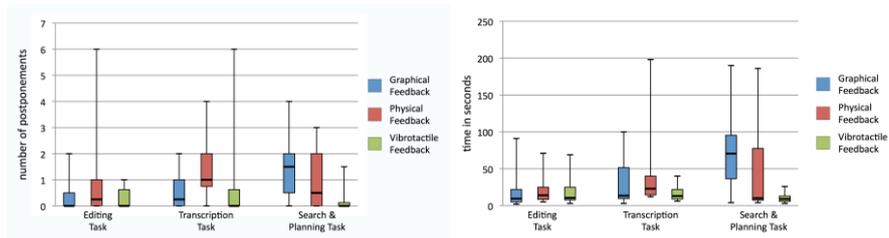
**Fig. 6.** *(left)* Boxplot of the number of trainings postponed for each task *(right)* and of the timespan until the first training session was started after the initial alert.

Furthermore, regarding the time span until a training was started the results indicate that participants took significantly longer to start with the training under the graphical feedback condition ($M = 42.22$, $SD = 28.54$) than under the vibrotactile condition ($M = 16.58$, $SD = 9.71$).

Analyzing the times to return to the main task after a training session, we found surprisingly short time spans. Again we found a main effect for the type of feedback, indicating that the time to return to the main task was longer under the graphical feedback condition ($M = 6.42$, $SD = 1.65$) than under the vibrotactile condition ($M = 4.79$, $SD = 1.24$) or the physical feedback condition ($M = 3.97$, $SD = 1.71$).

We counted the number of times a user looked at the taskbar feedback icon in the digital avatar situation from the Tobii screen recordings (cf. **Fig. 7**). We realized that participants paid attention to the feedback, although some stated that they were not aware of the feedback (Task 1: $M = 3.41$, $SD = 3.34$; Task 2: $M = 3.58$, $SD = 1.56$; Task 3: $M = 3.92$, $SD = 3.29$). This discrepancy between their evaluation and the collected data might be caused by the fact that they frequently glanced at the icon when switching between two applications (word and a browser, for example). In this situation, they might not have intentionally looked at the icon, but they nevertheless shortly checked the state of the graphical feedback.
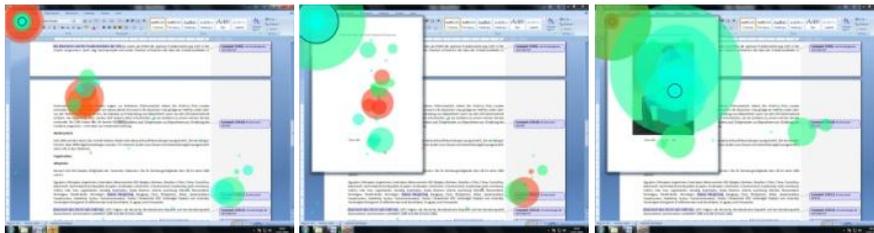


**Fig. 7.** Snapshot of a participant, who is looking at the physical avatar before starting the training. First, the user is working on the primary task (*left*). The shaking avatar attracts his attention (*center*) and s/he starts the training (*right*).

### 5.2 Emotional State Measurements Results

The results of the NASA-TLX questionnaires are depicted in Table 2 and Table 3.

| TLX Value | Effect | *F* | *p* |
|-----------|--------|-----|-----|

| | | | |
|---|---|---|---|
| Mental Demand | Task | $F_{2,22} = 3.260$ | 0.058 |
| | Feedback | $F_{2,22} = 1.133$ | 0.340 |
| Temporal Demand | Task | $F_{2,22} = 14.225$ | < 0.001* |
| | Feedback | $F_{2,22} = 1.000$ | 0.384 |
| Performance | Task | $F_{2,22} = 25.379$ | < 0.001* |
| | Feedback | $F_{2,22} = 2.363$ | 0.118 |
| Effort | Task | $F_{2,22} = 7.493$ | 0.010* |
| | Feedback | $F_{2,22} = 2.401$ | 0.114 |
| Frustration | Task | $F_{2,22} = 7.503$ | 0.011* |
| | Feedback | $F_{2,22} = 2.684$ | 0.091 |
| Workflow | Task | $F_{2,22} = 4.395$ | 0.049* |
| | Feedback | $F_{2,22} = 5.482$ | 0.012* |
| Disturbance due to Feedback | Task | $F_{2,22} = 0.409$ | 0.569 |
| | Feedback | $F_{2,22} = 20.425$ | < 0.001* |
| Disturbance due to Training | Task | $F_{2,22} = 4,515$ | 0.023* |
| | Feedback | $F_{2,22} = 6,252$ | 0.007* |

Table 2: Main effects for TLX dimensions. Starred results are significant for $\alpha = 0.05$.

A two-way within-subjects analysis of variance (ANOVA) was conducted to evaluate the effect of feedback strategies and task type on the various TLX measures. The dependent variables were continuous TLX ratings of 0 (low) to 20 (high). Results are summarized in Table 2. There were no significant interaction effects. Post hoc analyses were conducted on the significant main effects. These consisted of paired-samples *t*-tests with familywise error rate controlled across the test using Holm's sequential Bonferroni approach. Significant differences between the means of pairs of conditions are presented in Table 3.

| TLX Value | Pair | *F* | *p* |
|---|---|---|---|
| Temporal Demand | Edit - Transcribe | $t(11) = -4.930$ | 0.000 |
| | Edit - Plan | $t(11) = -5.438$ | 0.000 |
| Performance | Edit - Plan | $t(11) = -7.910$ | 0.000 |
| | Transcribe - Plan | $t(11) = -3.893$ | 0.003 |
| Effort | Edit - Transcribe | $t(11) = -3.443$ | 0.005 |
| | Edit - Plan | $t(11) = -4.748$ | 0.001 |
| Frustration | Edit - Plan | $t(11) = -4.415$ | 0.001 |
| Workflow | Edit - Transcribe | $t(11) = -3.273$ | 0.007 |
| Workflow | Graphical – Physical | $t(11) = 3.785$ | 0.003 |
| Disturbance due to Feedback | Graphical - Vibration | $t(11) = -4.899$ | 0.000 |
| | Physical - Vibration | $t(11) = -4.529$ | 0.001 |
| Disturbance due to Training | Graphical – Physical | $t(11) = 4.597$ | 0.001 |

Table 3. Significant mean differences along TLX dimensions between pairs of conditions.

### 5.2.1 Type of Task

The type of task had a significant main effect on reported temporal demand, performance, effort and frustration (cp. **Fig. 8**).
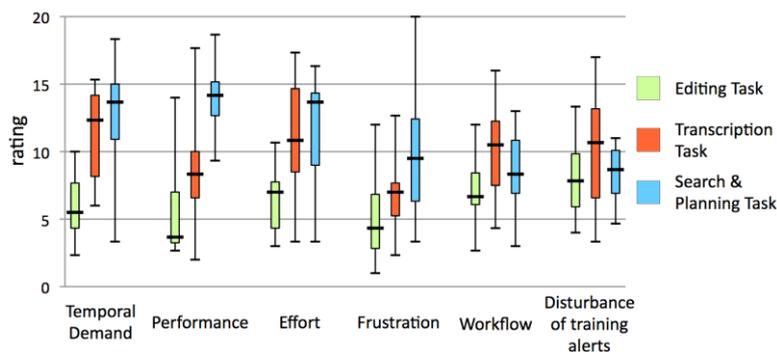


**Fig. 8.** Boxplot of average NASA-TLX scores for the 3 task types (0 = low, 20 = high).

The task load was assessed as lowest for the editing task, while the planning task was assessed as the most demanding. Besides this the type of task also had a significant main effect on the workflow and the disturbance due to training alerts. The impact on the workflow and the disturbance due to training alerts was rated as most severe for the transcription task.

### 5.2.2 Type of Feedback

The type of feedback had a significant main effect on participants' reported interruption of the workflow (cf. **Fig. 9**). The physical avatar was rated as less disruptive to the workflow as the other two feedback modalities.
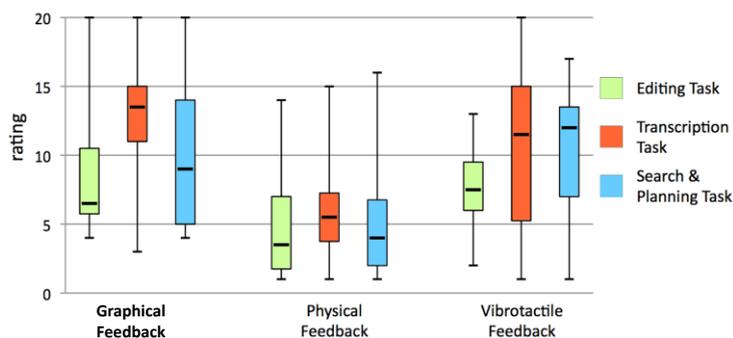
**Fig. 9.** Boxplot depicting participants ratings of the Impact on Workflow (0 = low, 20 = high).

The type of feedback also had significant main effects on the perceived disturbance due to both continuous feedback on the sitting position as well as the alert for a training session, even though the patterns are different (cf. **Fig. 10**). While participants rated the vibrotactile feedback as more disturbing than the digital and physical feedback, digital and vibrotactile feedback were assessed as more disturbing than the physical feedback when providing alerts for training.
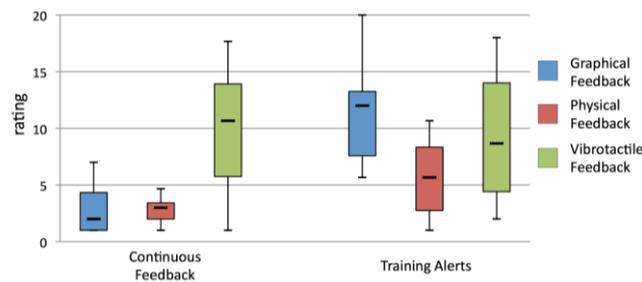


**Fig. 10.** Boxplot of the perceived disturbance due to continuous feedback and alerts (0 = low, 20 = high).

## 6    DISCUSSION

The fact that the vibrotactile feedback resulted in quite low response times across all three task types (cf. **Fig. 6**, *right*) is in line with participants' comments that they intuitively tried to stop the continuous vibration feedback highly soon, because it was annoying for them. Four participants reported during the survey that the vibrotactile feedback has been "*extremely disruptive*" during the task and they "*might switch it off if they had to use it in a long-term study*". The vibrotactile feedback was harder to ignore for them compared to the graphical and physical feedback. In contrast to the graphical and physical feedback, the vibrotactile feedback was clearly noticeable with the lightest feedback status already (30% of the full vibration strength). With increasing strength (up to 80%), the feedback's disrupting effect was also incrementing. The most obvious way to deal with the feedback was to simply react to the alert and start with the training.

**Fig. 11.** During the training session (1), the participant checks the current time in the lower right corner (2) and already plans an actions in the primary task window (3). At the end of the training, the attention is focused to the close button (4). Immediately after closing the training window (5), the participant clicks the browser tab as planned before (6).

The general short times we got for returning to the main task seem to be caused by the type of interrupt we are dealing with; since there is low cognitive load during the physical training and the participants are informed about how long the task lasts (through the countdown timer), they already plan the next steps of the primary task. **Fig. 11** depicts the tracking data of one participant, where she looked at the browser's tab 19 s before the end of the training. After the training was finished, it took her only 2 s to click exactly this tab. This phenomenon was highly interesting and it has been observed with other 6 participants. Although the tendency to postpone a training appeared to be lowest under the vibrotactile feedback condition, most of the participants had the impression that the vibrotactile feedback was assessed as the most disturbing one, which would result in a termination of the system after a short period of use.

In the participants' feedback comments, we noticed the same comments saying "I noticed the vibrotactile really fast – but it was so disturbing – I just wanted to turn it off". Similar to McFarlane's conclusion, we found that giving people the control when to react to an interrupt might cause the side-effect that people always try to postpone interrupting alerts [17].

Finally, in our study, we also noticed that 7 of the participants did not notice the shaking physical avatar at the very beginning. But they also mentioned that once they noticed it, it was less disturbing, because it was not in their field of view.

## 7 CONCLUSION AND FUTURE WORK

Summarizing it turns out that the proposed posture chair with the physical (ambient) interrupt motivates people to improve their sitting behavior – even if they have to work focused on a primary task. The comments and the data from our study also demonstrated that an additional (visual) feedback will be accepted – especially if it does not interfere with the working screen. The results of the first study motivate us to improve the current system. Moreover, we plan to do a long-term field study using 12 chairs for a period of three months. This study will be done in cooperation with physical therapists with the overall goal to demonstrate the benefits of a posture chair setup.

## ACKNOWLEDMENTS

## REFERENCES

1. Adamczyk, P. D. and Bailey, B. P. 2004. If not now, when? The effects of interruption at different moments within task execution. In *Proceedings of CHI '04*. ACM, New York, NY, 271-278.
2. Arroyo E. and Selker T., 2003. Arbitrating multimodal outputs: Using ambient displays as interruptions, in Human-Computer Interaction: Theory and Practice (Part II) - Proceedings of HCI International 2003, Vol. 2, 591-595.
3. Arroyo, E., Selker, T., Stouffs, A., 2002, Interruptions as Multimodal Outputs: Which are the Less Disruptive? Multimodal Interfaces, IEEE International Conference on Multimodal Interfaces, p. 479, 2002.
4. Bailey, B. P., Konstan, J. A, and Carlis, J. V., 2001. The effects of interruptions on task performance, annoyance, and anxiety in the user interface. INTERACT 2001, 593-601.
5. Bailey, B.P., Konstan, J. A., & Carlis, J.V., 2000. Adjusting windows: Balancing information awareness with intrusion. In P. Kortum & E. Kunzinger (Ed.), Proceedings of the 6th conference on human factors and the Web: Doing business on the Web. Austin TX.
6. Beach T.A., Parkinson R.J., Stothart J.P. and Callaghan J.P., 2005. Effects of prolonged sitting on the passive flexion stiffness of the in vivo lumbar spine, Spine J. 5 (2005), pp. 145-154.
7. Czerwinski, M., Cutrell, E. and Horvitz, E. Instant Messaging: Effects of Relevance and Timing. In People and Computers XIV: *Proceedings of HCI 2000*, 71-76.
8. Czerwinski, M., Horvitz, E., and Wilhite, S. 2004. A diary study of task switching and interruptions. In *Proceedings CHI '04*. ACM, New York, NY, 175-182.
9. Daian, I., van Ruiten, A. M., Visser, A., and Zubic, S. 2007. Sensitive chair: a force sensing chair with multimodal real-time feedback via agent. In Proceedings of the 14th European Conference on Cognitive Ergonomics: invent! Explore! vol. 250. ACM, 163-166.
10. Ertel, M.; Junghanns, G.; Pech, E.; Ullsperger, P. 1997. Effects of VDU-assisted work on health and well-being. Research Report 762, Federal Institute for Occupational Safety and Health (BAuA).
11. Hart, S.G., & Staveland, L.E. 1988. Development of a NASA-TLX (Task load index): Results of empirical and theoretical research. In P.A. Hancock and N. Meshkati (Eds.), Human Mental Workload, 139-183.
12. Horvitz, E. and Apacible, J. 2003. Learning and reasoning about interruption. In *Proceedings of the 5th international Conference on Multimodal interfaces* ACM, New York, NY, 20-27.
13. Iqbal, S. T. and Bailey, B. P. 2008. Effects of intelligent notification management on users and their tasks. In *Proceeding of CHI '08*. ACM, NY, 93-102.
14. Iqbal, S. T. and Horvitz, E. 2007. Disruption and recovery of computing tasks: field study, analysis, and directions. In *Proceedings of CHI '07*. ACM, NY, 677-686.
15. Kingma I., van Dieen J.H., 2009, Static and dynamic postural loadings during computer work in females: Sitting on an office chair versus sitting on an exercise ball, Applied Ergonomics, Volume 40, Issue 2, pp. 199-205
16. Lieberman, H. 1997. Autonomous interface agents. In *Proceedings of CHI '97*. ACM, New York, NY, 67-74.
17. McFarlane, D., 1999. Coordinating the interruption of people in human-computer interaction. Proceedings of Interact '99, pp. 295-303.
18. Pfendler, C., 1991. Vergleichende Bewertung der NASA-TLX-Skala bei der Erfassung von Lernprozessen. Wachtberg: Forschungsinstitut für Anthropotechnik, Bericht No. 2.

19. Rivera, D. 2005. The effect of content customization on learnability and perceived workload. In *Proceedings of CHI '05,* ACM, New York, NY, 1749-1752.
20. Salvucci, D. D. and Bogunovich, P. 2010. Multitasking and monotasking: the effects of mental workload on deferred task interruptions. In *Proceedings of CHI '10*. ACM, New York, NY, 85-88.
21. Wickens, C.D., Hollands, J. G. Engineering Psychology and Human Performance. Harper Collins, 1992.
22. Zheng, Z., Morrell, J.B. A. vibrotactile feedback approach to posture guidance, Haptics Symposium, 2010 IEEE pp. 351-358.