Situational impairment due to walking with conversational versus graphical interfaces

Georgios Marentakis*
Arman Balic
georgios.marentakis@hiof.no
arman.balic@hiof.no
Department of Computer Science and Communication
Østfold University College
Halden, Norway

ABSTRACT

We present a study that investigates situationally-induced impairments that appear while walking and interacting with a smartphone using a conversational versus a graphical user interface. In a controlled experiment, participants performed a mixture of tasks under conditions that manipulated Mobility (standing or walking) and interaction Modality (graphical or conversational) while we measured walking performance, task performance, and subjective workload. Although walking while interacting did not impair task performance significantly in either case, it significantly increased perceived workload when interacting with the GUI but not when interacting with the CUI. Furthermore, compared to a control walking-only condition, walking performance deteriorated less with conversational than with graphical interaction. Finally, interaction with the conversational interface was slower than this with the graphical. The results testify to an increased potential for conversational interfaces to support walking interactions but also show that due to technology limitations this does not manifest in a task performance advantage yet.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; **Auditory feedback**; **Empirical studies in ubiquitous and mobile computing**.

KEYWORDS

situational impairment, walking, conversational user interfaces

ACM Reference Format:
Georgios Marentakis and Arman Balic. 2024. Situational impairment due to walking with conversational versus graphical interfaces. In *Audio Mostly 2024 - Explorations in Sonic Cultures (AM '24), September 18–20, 2024, Milan, Italy*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3678299.3678307

AM '24, September 18–20, 2024, Milan, Italy

ACM ISBN 979-8-4007-0968-5/24/09
https://doi.org/10.1145/3678299.3678307

1 INTRODUCTION

Graphical User Interfaces (GUIs) supporting direct manipulation provided a significant breakthrough for Human Computer Interaction and have been the dominating interaction paradigm. GUIs make full use of our visual system and by coupling it to manual interactions create a very efficient interaction paradigm. However, when vision is occupied with another task, usability problems emerge which make interacting with a GUI difficult and even dangerous.

The term *situationally-induced impairments and disabilities* (SIIDs) [56] was introduced to describe situations in which the usability of a computer system is affected by context of use. Several cases of (SIIDs) have been reported in mobile interaction use cases due to walking, variations in the environmental conditions, or encumbrance [3, 46, 50, 51, 58, 61, 66]. Efforts have been made to counteract their impact by design [13, 14, 26, 36, 56].

Sound has often been used to provide feedback in graphical user interfaces. Sound in graphical user interfaces enhances user experience but also usability especially when using small screens or when visual attention is occupied e.g., [7, 45]. Prominent approaches to designing sound for the user interface are auditory icons [12], earcons [6, 10], spearcons [62], and sonification [20]. These have also been used to create exclusively auditory interactions. They provide usable alternatives for presenting information whose practical significance increases considerably in situations in which the visual modality is occupied or impaired. Auditory icons are generally considered to lead to faster learning and quicker responses as long as they are supported by an appropriate metaphor [11].

Even if speech feedback is very important for visually impaired users [48], it has received less attention within the auditory display community, as it may result in potentially long and repetitive messages. Some of these aspects are addressed in spearcons, speeded speech messages, which combine the learnability of speech with good user performance [62].

The aforementioned approaches to auditory display are largely derived from the dominant direct manipulation interaction paradigm. However, recent advances in speech recognition, natural language processing, and generative language models helped improve usability issues associated with conversational user interfaces (CUIs) whose popularity is steadily increasing and as a result receive increased attention in recent years. It is therefore interesting to investigate how they compare to visual direct manipulation interfaces in situations in which the visual modality is occupied and situational impairments emerge. Driving or walking are good examples of such use cases [30, 50, 51].

Interacting in contexts which constrain the resources available for interaction, as for example when walking or driving, may also be considered as a dual-task. In dual-tasks, a performance impairment is typically expected that may be observed on one or both tasks. In this sense, interacting while walking or driving may lead to an impairment to both interaction performance (which has been considered to be a situational impairment) but also on walking or driving performance. For example, fragmentation of attention has been observed when asking participants to perform browsing tasks while moving through urban environments due to attending to the challenges involved [43, 44]. Furthermore, both walking and task performance are impaired when interacting with a display while walking [5, 34, 45].

The emerging dual-task cost is, however, smaller when tasks are not in the same modality [45, 64]. In special cases, it may even disappear if no resource limitation to the secondary task is posed by the first task. As interaction with CUIs is done primarily in the audio modality, which interferes less with visual or spatial tasks, the amount of situational impairment due to walking should be smaller when using a CUI compared to when using a GUI. Such predictions have been verified when driving [19, 25, 35, 37], for example.

However, walking is different than driving as the whole body is involved. Even if interaction with a CUI is done in the audio modality, the cognitive resources required are not negligible. Furthermore, despite the progress in speech processing usability problems still exist and are not uncommon. It is therefore interesting to investigate both whether the situational impairment due to walking will be reduced when interacting with a CUI in comparison to a GUI but also whether walking performance is affected less by conversational compared to graphical interactions. To investigate these two aspects, we designed and performed an experiment comparing interaction with a graphical and a conversational interface in a standing and a walking scenario. We proceed by presenting the background, the experiment, results, and discussion.

2 BACKGROUND

SIIDs may emerge due to changes in the context and the user environment [24, 55, 56, 66, 67] but also as the result of fragmentation in attention [43, 57, 63]. Walking, changes in light [2, 29, 31] or temperature [15, 49, 52], or encumbrance [40–42] have been identified as sources of situational impairments for mobile interaction with smartphones.

Walking, which is the focus of this article, may affect task performance and perceived mental workload when interacting with a smartphone using a graphical interface [18, 39]. Walking impacts target selection accuracy, reduces walking speed, and increases perceived workload [30]. Walking degrades visual performance when reading and when accomplishing visual search tasks [38]. For a reading task, [54] reports that walking increased error rate by 24%, target selection time by 31%, cognitive load by 16%, and reading speed by 19%. Furthermore, walking results in changes in ambient light, ambient noise, encumbrance and temperature which also affect mobile interaction [28].

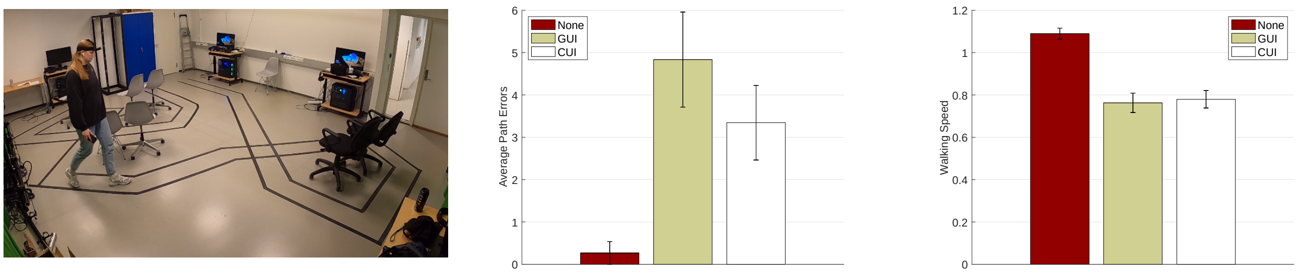
Users compensate for disturbances due to walking by reducing walking speed; [5] report that users reduced walking speed to maintain good target acquisition performance when mobile. Target size can also be adjusted to help walking users. Task completion time improved when designing a walking user interface that increases target size when users move [26, 68]. NoShake [47] shifts the screen in the opposite disturbance direction which results in a user experience improvement. WalkType [13] utilizes the built-in accelerometer to reduce error rates while walking by 45.2% and improved typing speed by 12.9% [13]. Tracing is a more effective method compared to tapping or handwriting for text entry when using a smartphone when mobile [60] and several alternative one-hand or gesture controlled interaction methods [27, 32, 53] or back of device methods [4] have been proposed (but not evaluated in mobile settings).

Audio feedback can compensate for the disturbances from walking [8] as users need not look at the screen. Gesture input coupled with audio feedback has been shown to significantly improve mobile interaction [9, 34, 45, 69] and audio can be delivered using different reproduction techniques [33]. Multiple resource theory accounts for improvements due to 'eyes-free interaction' [63–65]. When two tasks demand the same level of a given perceptual, cognitive or motor dimension, they will interfere with one another and performance will be negatively affected [65]. Since smartphone interaction when walking requires dividing visual attention, dual-task interference may be reduced by enabling interaction in the auditory channel.

Conversational User Interfaces enable interaction by talking to a computer system [21, 23]. In a recent survey [23], text entry and typing, application control, speech analysis, conversational agents, spoken output, & probes were found as common application themes for CUIs. Munger et al [37] found that CUIs provide a significant advantage over a touch interface in terms of distraction and length of interaction while driving. He et al. [19] found that handheld texting increased the brake response time, among other safety-related factors, in comparison to the CUI use. CUIs on smartphones enabled the performance of several tasks such as placing calls, destination entry [37], controlling entertainment [16], reading and transcribing messages, or similar tasks without demanding the visual or physical attention of the driver [25, 35]. Still, complex CUI interactions show significant attention demands [37] and may significantly impair driving [19, 22].

2.1 Summary and Research Questions

While CUIs have received significant attention, the extent to which the language based auditory interactions they offer can reduce SIIDs due to walking and reduce the impact on users' walking performance has not been investigated. Research on using CUIs while driving provides encouraging results, however, it also shows that using CUIs comes with significant cognitive demands which require a detailed and specific investigation. The research question we investigate is therefore whether CUIs can reduce SIIDs due to walking and lead to improved walking performance when performing common smartphone tasks in comparison to touch-based GUI interaction.



(a) (b) (c)

Figure 1: a. a participant walking along the path, b. average number of path errors, and c. average walking speed. Mean and standard error are shown.

3 EXPERIMENT

The experiment we designed followed an interaction Modality (GUI, CUI) × Mobility (standing, walking) within-subjects factorial design. Factor interaction Modality enabled comparing between interacting with a GUI versus a CUI and factor Mobility comparing between interacting while standing and walking. In the four conditions of the experiment, participants performed the same number of a mixture of common texting and selection smartphone tasks while standing and while walking on a pre-defined route. There was also a control condition in which participants walked without interacting with the smartphone. Hypotheses were formed based on the background and examined on the basis of walking performance, task performance, and workload estimated using the NASA Task Load Index (NASA-TLX) [17].

3.1 Apparatus and Materials

A *walking path* was marked on a lab room floor which contained obstacles and included loops and sharp turns (Figure 1a). Path width was 30cm and length was 30 meters. Participants wore a pair of AirPods (2nd generation) and used an Apple iPhone 12 Pro with iOS 15.2 and Siri set to participant's native language which was the same for all participants. Siri was reset after each trial of the experiment. Participants held the iPhone in their dominant hand. They called "Hey Siri" to activate CUI interaction. They wore a GoPro Hero 8 camera which recorded the experiment. The phone screen was also recorded.

3.2 Participants

Sixteen participants took part. Participants were screened for speech or motor impairments and one withdrew due to having such an impairment. Age was between 18 to 50 years, ($\mu = 29.8, s = 6.75$), 9 male and 6 female. Participants provided informed consent.

3.3 Tasks

A mixture of tasks was administered to the participants consisting of both text entry (*texting*) and (*selection*) tasks which were subsequently pooled in order to obtain a representative dataset for analysis. Participants were not allowed to use the GUI when engaging with the CUI and vice-versa. Tasks were selected based on statistics of common smartphone tasks [1]. The experimenter demonstrated how tasks can be performed using the CUI or the GUI and made sure participants could repeat the steps before engaging in the experiment.

Texting tasks required participants to respond to blocks of four short questions each. The intention was to simulate a short conversation. In each occasion, they received a question in an SMS to which they replied. Questions were selected randomly from a list of 15 possible ones, for example: what was the last city that you visited?, what is your favourite thing to eat?, what city were you born in?, etc. When using the CUI participants used Siri to listen and dictate a reply to the SMS. They said 'Repeat' to hear the message again if they wanted to. When using the GUI, participants read the incoming message and simply typed and sent the response. Participants were instructed to provide their response in one text message and to respond in a natural way as if responding to a friend. No other constraints were given. A new text message was sent a few seconds after a reply was received.

There were three different selection tasks which required a number of steps so that they were completed. The first task was *planning a trip home* (via the Maps app) and involved the following steps: find out how long it is to get home, find out what time the next bus goes, set a reminder to leave, and find the closest bus stop. The second task was *music selection* and involved the following steps: play a specific song from the phone music collection, listen for 15-20 seconds, increase the sound level, select the next song, listen for another 15-20 seconds, then select another song. The third task was *planning an outdoor hiking trip* with a friend and involved the following steps: check the weather for the next day, set a reminder to wake up and dress well, and send a message to the friend to confirm. Participants used either the smartphone GUI or CUI to perform the tasks.

3.4 Procedure

Participants provided informed consent and received information about the experiment process, tasks and conditions, and the NASA-TLX questionnaire. They then filled a short *pre-experiment questionnaire* assessing prior experience with smartphones and CUIs. Then participants walked along the path a few times (2-3) to familiarize themselves. Natural walking speed was then recorded by asking participants to walk on the path first in a clockwise direction and then in a counter-clockwise direction at a comfortable walking pace.

Subsequently, participants engaged in the experiment conditions (GUI standing, GUI walking, CUI standing, and CUI walking). All tasks described in Section 3.3 were performed by each participant. The order of presentation of Mobility and Modality conditions was counterbalanced using a Latin Square.

In each Mobility and Modality combination, there were two trial blocks: one for the texting and for the selection tasks. In the texting block, participants responded to three blocks of four randomly selected text messages. In the selection block, participants performed the three selection tasks in a randomized order. Before engaging in each condition, the way tasks could be performed in each configuration was demonstrated by the experimenter. Subsequently, participants performed two training trials during which they received feedback from the experimenter who made sure they could perform the tasks as instructed.

After completing each condition, participants filled a NASA-TLX questionnaire. Finally, participants shared comments and thoughts about their interaction experience. The experiment took between 45 and 65 minutes including a 15 minute break, if they felt tired. Figure 1a shows a participant performing the experiment.

3.5 Hypotheses

The literature indicates that a degradation in walking performance can be expected [5] when walking and performing the experiment tasks simultaneously. Therefore **(H1)** walking speed and number of path following errors will deteriorate while interacting with the GUI and with the CUI. Interference between walking and interacting will also affect task performance. It follows that **(H2)** task completion speed and accuracy will deteriorate and subjective workload will increase when interacting while walking compared to interacting while standing both when interacting with the GUI and with the CUI. However, given that interacting with a CUI interferes less with walking compared to interacting with a GUI, it is reasonable [63, 64] to expect that that **(H3)** significantly fewer path errors and faster walking speed is expected when interacting with a CUI compared to when interacting with the GUI. Finally, less interference between the walking and the interaction task will also assist task performance. It follows that **(H4)** fewer task errors, shorter task completion times, and a lower workload are expected when using a CUI than when using a GUI while walking.

4 RESULTS

In the trials, each participant performed the selection tasks and responded to the text message blocks (see Section 3.3) in the four conditions of the experiment: standing and GUI interaction, standing and CUI interaction, walking GUI interaction and standing and CUI interaction. Results were pooled over the performed tasks for analysis.

As mentioned above, all trials were video taped using a 3rd person and a 1st person camera view. Then they were analyzed manually after the experiment was finished and coded by the experimenters. Trial completion time, trial completion errors, and path following errors were marked in the videos. Trial completion time was estimated as the time difference between the beginning and end of each trial time points as these were marked in the videos. The steps performed to accomplish the tasks by each participant were also analyzed and compared to the optimal sequence of actions that was necessary to complete the task as this has been defined by the experimenters and communicated to the participants. Situations in which the user deviated from the optimal sequence due to own action or due to an unexpected system response were classified as task errors. Other errors and mistakes were also coded. The rate of such instances per user and trial (error rate) was used to perform the statistical analysis in terms of task accuracy.

Walking speed was calculated by counting the number of laps converting this to meters and then dividing by the walking duration to yield a number of meters per second. This was done both when interacting and when walking without interacting. Furthermore, the number of path deviation errors (or walking errors), defined as instances of deviations from walking with the preferred speed inside the walking path in each trial in the video was coded. The average number of such walking errors per trial and participant was used in the analysis.

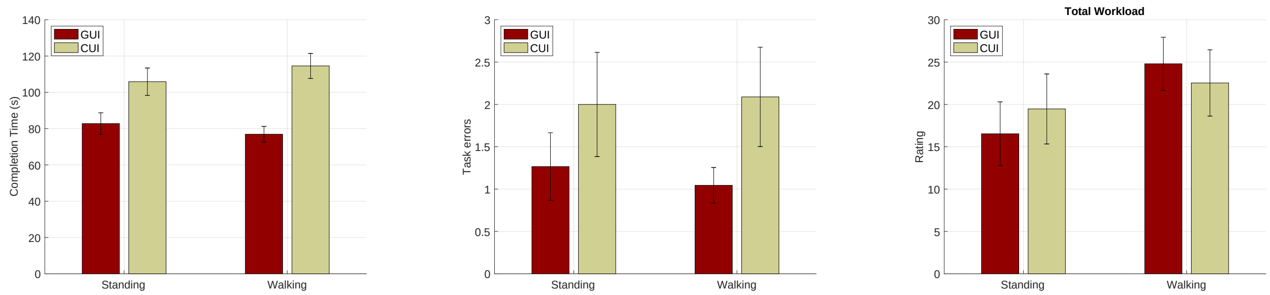
The types and frequencies of path following and task errors that were identified and counted are described in Section 4.6. Coding and video segmentation was done in MaxQDA while the statistical analysis was done in R. P-values and effect sizes are provided as appropriate. The generalized eta squared (η_p^2) is used in case of ANOVAs and Kendall's W and rank-biserial coefficient (r) for Friedman and Wilcoxon tests respectively [59].

4.1 Pre-Experiment Questionnaire

All participants were iPhone users, used Apple's Siri as their main voice user interface, and owned a smartphone for longer than 5 years. 16.7% reported having used CUIs often, 41,7% on occasion, 16.7% once or twice, and 25% never. Furthermore, 16.7% of the participants reported using Apple's Siri and Google Assistant actively. CUIs were used when driving or at home, primarily for sending and responding to text messages, controlling music players, and conducting web searches. Less frequent tasks included setting reminders, asking for directions and locations, and checking the weather. 50% have faced difficulties when using a voice assistant for mobile interaction mostly speech recognition errors and misunderstandings that led to CUIs issuing no or wrong commands. Only 16,7% reported speech recognition issues when doing text entry, sending text messages, and writing notes. When asked to rate their experience with using a voice assistant from 0 to 7, the average score was 3.9 (median score of 4.5).

4.2 Walking Performance

Figure 1b shows the average path errors and Figure 1c the average walking speed in the experiment conditions. *Walking speed* data were distributed normally as verified by visual inspection and using a Shapiro-Wilk normality test. A one-way repeated measures ANOVA with interaction Modality as independent variable was performed on walking speed. Interaction Modality had three levels



(a) Task completion time **(b) Number of task errors** **(c) Total Task Workload**

Figure 2: Average number of task completion errors, task completion time, and total workload for the Modality and Mobility conditions in the experiment pooled over the exerimental tasks.

(GUI, CUI, NONE) with NONE denoting walking without performing a task. The main effect of interaction Modality on walking speed was significant, $F(2, 28) = 57.77$, $p<0.001$, $\eta_p{}^2=0.53$. Pairwise comparisons with Holm correction showed that both when interacting with the CUI and with the GUI walking speed was slower than normal ($p<0.001$). Walking speed did not change significantly when interacting with the GUI compared to the CUI. *Path Errors* per participant data were not distributed normally as verified by visual inspection and using a Shapiro-Wilk normality test. A Friedman test was performed with Modality as independent variable as above. The main effect of Modality was significant $\chi^2=2.5$, $p<0.001$, $W = 0.9$. Wilcoxon sign rank tests showed that the number of path errors was significantly higher when using the GUI compared to when using the CUI ($p=0.0018$). Furthermore, the number of path errors was significantly higher compared to walking without interacting for both interaction Modalities ($p<0.001$).

4.3 Task Completion Time

Average completion times can be seen in Figure 2. Data were distributed normally after log transformation as shown by visual inspection and using a Shapiro-Wilk test. A two-way Modality (GUI, CUI) × Mobility (Standing, Walking) ANOVA was performed on log-task completion time. The effect of Modality was significant, $F(1,14) = 32.48$, $p<0.001$, $\eta_p{}^2=0.43$. Task completion was faster with the GUI. No other effects or interactions were significant.

4.4 Task Errors

Average number of task errors can be seen in Figure 2. Data were not distributed normally as shown by visual inspection and using a Shapiro-Wilk test. A Wilcoxon rank sum test with continuity correction was used to compare the number of task errors when using the GUI and when using the CUI in the standing and walking conditions. The effect of Modality was not significant neither when standing nor when walking. The effect of Mobility was not significant neither when using the GUI nor the CUI.

4.5 Workload

Mental demand, physical demand, temporal demand, perceived performance, effort and frustration can be seen in Figure 3. Data were not normally distributed. Total workload has been presented in Figure 2c. The effect of Mobility was examined by applying a Wilcoxon test to the data from the trials corresponding to each Modality level (GUI or CUI) while the effect of Modality was examined by applying a Wilcoxon test to the data from the trials corresponding to each Mobility level (standing or mobile).

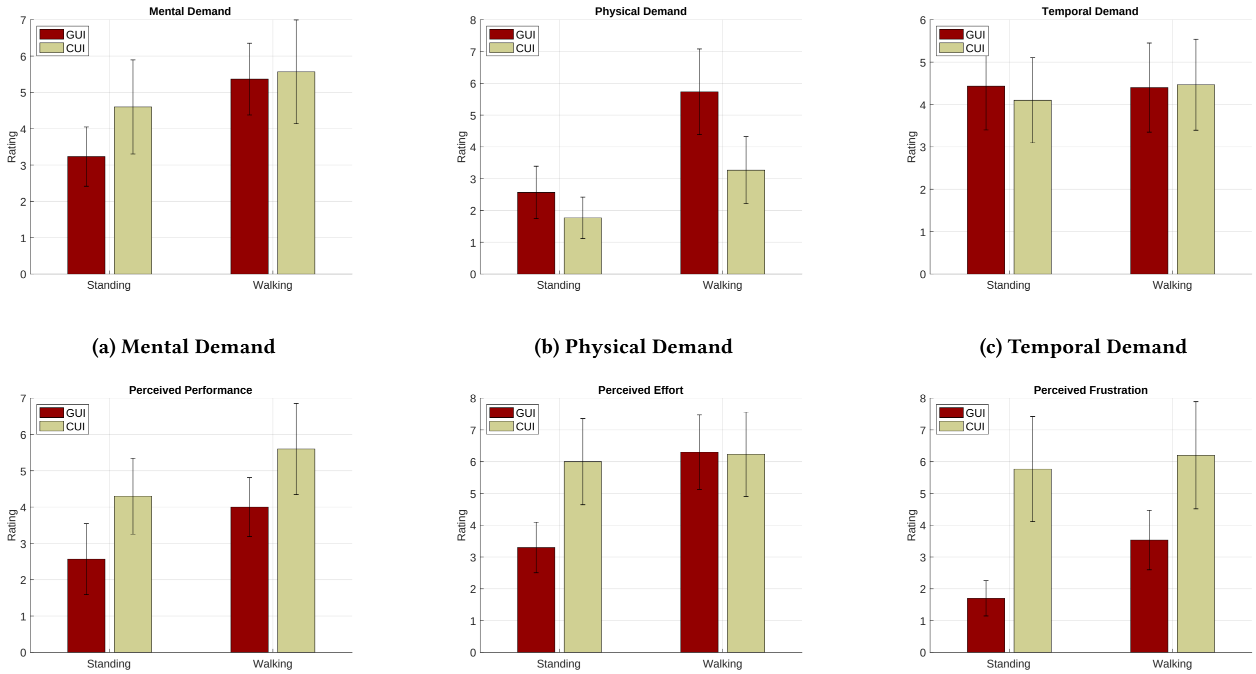
The effect of *Mobility* was significant for *total workload* ($p=0.01$, $r = 0.31$), *mental demand* ($p=0.02$, $r=0.28$), *physical demand* ($p=0.004$, $r=0.37$), *perceived performance* ($p=0.04$, $r=0.25$), *perceived effort* ($p=0.004$, $r=0.36$) and *perceived frustration* ($p=0.02$, $r=0.28$) when using the GUI but not when using the CUI. *Total workload*, *mental demand*, *physical demand*, *perceived effort*, and *perceived frustration* increased when walking and using the GUI but not when using the CUI. *Perceived performance* decreased significantly also when using the CUI ($p=0.04$, $r=0.25$). Both when using the GUI and when using the CUI, participants felt they performed worse when walking. The effect of *Mobility* was not significant for *temporal demand* for both interaction Modalities.

The effect of *Modality* was significant when walking for *physical demand* ($p=0.017$, $r=0.3$) but not when standing. When walking, *physical demand* was higher when interacting with the GUI compared to the CUI. The effect of *Modality* was significant for *perceived effort* ($p=0.03$, $r=0.27$) and *perceived frustration* ($p=0.04$, $r=0.26$) when standing but not when walking. Both were significantly higher when using the CUI compared to the GUI. The effect of *Modality* was not significant for *total workload*, *mental demand*, *temporal demand*, and *perceived performance* both when standing and when walking. We note that the effect of *Modality* on perceived performance was marginally significant as performance was worse on average when using the CUI compared to the GUI ($p=0.0516$, $r=0.25$).

4.6 Walking and Task Error Types

To gain further insight task errors and walking errors were classified based on the observations in the video on the MaxQDA analysis software.

4.6.1 Path Errors. The following walking errors were observed: stepping on or over the line, readjusting walking path, walking outside the outlined path, failing to follow the path (i.e. path related),



(a) Mental Demand (b) Physical Demand (c) Temporal Demand

(d) Performance (e) Effort (f) Frustration

Figure 3: Average ratings of NASA-TLX workload dimensions.

readjusting walking, slowing down, and stopping (i.e., walking related) and hitting an obstacle (i.e., obstacle related). Figure 4a shows the average path errors per trial in the conditions tested in the experiment. Path errors were most frequent (especially stepping on or over the line outlining the path), followed by adjustments in the walking pattern, and then hitting obstacles. It is also visible that interaction with the CUI leads to fewer walking errors on average and that interference with the walking pattern is minimal.

4.6.2 Task Errors. Different task errors were observed when interacting with the GUI (Figure 4b) compared to when interacting with the CUI (Figure 4c). When interacting with the GUI, some errors were related to the steps required to complete the tasks as these have been considered for the experiments (step failure, missed step, extra step, resend message) while some were related to text entry (spelling error, illogical answer, double texting - replying to the same message twice). Engaging with the tasks while walking lead to an increase every type of error observed. Notably, failures to complete a step, taking extra steps, and spelling errors double while we also note misclicks which did not appear when standing.

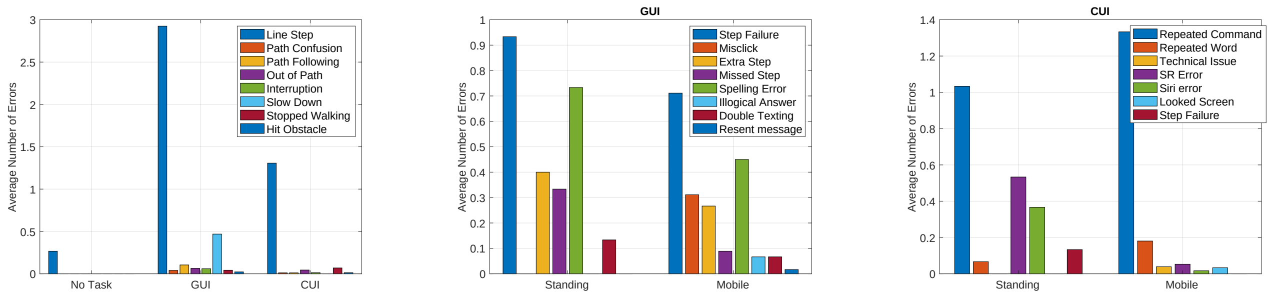
When interacting with the CUI, some step related errors reappear (step failure), however, we have several errors types that emerge due to interaction with the CUI (repeated commands or words, speech recognition or Siri error, technical issues e.g., dropouts) and cases in which participants looked at the screen. The impact of walking here is less evident and concentrated to interaction with the voice interface.

4.7 Summary of the results

We provide here a summary of the most important results: (1) walking performance deteriorated when interacting but less with the CUI compared to the GUI (fewer path errors), (2) interaction was significantly faster with the GUI but there was no difference in terms of task errors, (3) mental demand, physical demand, perceived effort, and perceived frustration, deteriorated due to walking when interacting with the GUI but not when interacting with the CUI, and (4) interacting with the GUI required higher physical demand when walking compared to the CUI. Even though perceived effort and perceived frustration were higher when interacting with the CUI when standing, the difference was not significant when walking.

5 DISCUSSION

Motivated by findings in dual-task performance, multimodal interaction, and situational impairment literature, we aimed to investigate whether the use of audio and speech for interaction, as in the CUI paradigm, can lead to an improved dual-task performance compared to a GUI when walking while interacting with a smartphone. We designed and performed an experiment in which participants engaged in texting and selection tasks and measured walking performance, task efficiency and effectiveness, and perceived workload to investigate this hypothesis. The significance level we used was 0.05 and the effect sizes we obtained were, as a rule, large or moderate [59].



(a) Walking Errors **(b) Task Errors (GUI)** **(c) Task Errors (CUI)**

Figure 4: Walking errors and task errors observed for both interaction modalities.

In line with existing literature [5, 18, 30, 38, 39, 43, 44, 54], and our Hypothesis 1, walking while interacting led to a deterioration in walking performance both in terms of path following errors and walking speed for both interaction techniques. However, in support to Hypothesis 3 this deterioration was significantly smaller when interacting with the CUI compared to when interacting with the GUI. This was evidenced by the significantly fewer path errors observed when performing the tasks using the smartphone CUI compared to the smartphone GUI. Observing walking error types (Figure 1b), in addition to more frequent line steps, slowing down also occurs more often when using a GUI than a CUI. It appears therefore that maintaining a relatively stable walking pattern is easier when interacting with a CUI compared to when interacting with a GUI. This result is consistent with studies that suggest that the auditory modality is particularly appropriate for interaction in mobile settings [9, 34, 45, 69] and with the predictions of multiple resource theory [63–65].

Hypothesis 2 predicting a deterioration in task performance with either interface due to walking was not verified as the effect of Mobility for task completion errors and task completion time was not significant. Since walking performance deteriorated due to interacting, it appears therefore that participants prioritised task over walking performance, which resulted in that the situational impairment due to walking in terms of task speed and accuracy observed here was not significant. Essentially, the dual-task cost was not attributed equally to the two tasks but was rather attributed to walking. More insight into how this was performed can be obtained by looking into the measured subjective workload.

The results indicate that maintaining a similar level of task performance in both standing and walking conditions was particularly demanding when interacting with the GUI but less so when interacting with the CUI. This is evidenced by the significant increase in perceived workload when interacting the GUI when mobile. Total workload, mental demand, physical demand, perceived effort, and perceived frustration increased significantly. Such a deterioration due to Mobility was only observed for frustration when interacting with the CUI. This is in support of Hypothesis 4, as it shows that in contrast to GUI interaction, subjective workload when interacting with a CUI is not affected by walking. Further support to this statement is provided by the significantly higher physical demand for GUI interaction compared to CUI interaction when walking.

Let us return here to our research question which investigated the extent to which CUIs can reduce situational impairment due to walking. Some of the evidence we provide is supportive. Even if task performance was not affected by walking for both interaction modalities, perceived workload does not increase due to walking when interacting with a CUI, while the opposite is the case when interacting with a GUI. This result supports the idea that the SIID due to walking is smaller for the CUI compared to GUI interaction. Turning our focus onto walking performance we see that even if walking performance deteriorated compared to preferred walking speed for both interaction paradigms, the deterioration was significantly smaller when interacting with a CUI compared to a GUI. Therefore, CUI interaction appears to support better users who are walking as they can walk closer to their normal walking speed.

On the other hand, despite the aforementioned advantages, interacting with the CUI was slower compared to the GUI and a comparable error rate was observed as there was no significant difference when comparing error rates in the two interaction modalities. It appears therefore that despite the smaller situational impairment and better support for walking, GUIs still provide an advantage in terms of interaction speed. On the one hand, delays associated with speech communication versus manual interaction could play a role here. These could originate in faster reading versus listening or cases in which more time was necessary in order to speak commands compared to selecting icons on screen. However, a significant part of the delay we observed is due to speech recognition and natural language processing errors. The overall higher frustration reported for the CUI can arguably be attributed to such events.

The classification of the task errors provides further insights. For both interaction paradigms, the frequency of several task error types increases when walking. For GUI interaction, these are errors related to the logical steps required to complete the task but also text entry errors and misclicks. For the CUI, however, it is mostly repeating commands or words to the system until these are performed in a satisfactory way, while other error types are not affected.

Other things being equal, it can be argued that the usability problems that appear with CUIs are likely to diminish as improvements in speech technology are deployed in modern systems. On the other hand, because the usability problems associated with touch-based interaction with GUIs originate in interference in the visual and motor components of interaction due to walking they are harder to eliminate. Given the significant pace of improvement of the non-mobility related CUI usability problems, it is perhaps quite likely that our findings supporting reduced SIID when using a CUIs for interaction while walking will be replicated and extended in the coming years.

6 CONCLUSION

We presented a study that investigated the extent to which Conversational User Interfaces can reduce situationally-induced impairment and provide better support to users interacting with smartphones while walking. In an experiment involving sets of texting and selection tasks performed while standing or walking, we measured walking performance, task performance, and perceived workload when interacting with a smartphone using a CUI or using a GUI. The results suggest that interacting with a CUI results in a smaller situationally-induced impairment as in contrast to a GUI, participants can maintain a similar level of performance when walking without an increase in subjective workload. Furthermore, participants walk closer to their normal walking speed while interacting with a CUI compared to a GUI. However, usability problems related to language processing result in that tasks are still faster to complete with GUIs compared to CUIs.

REFERENCES

[1] [n. d.]. Smart speakers: use case frequency U.S. 2020. https://www.statista.com/statistics/994696/united-states-smart-speaker-use-case-frequency/

[2] Leon Barnard, Ji Soo Yi, Julie A. Jacko, and Andrew Sears. 2005. An empirical comparison of use-in-motion evaluation scenarios for mobile computing devices. *International Journal of Human-Computer Studies* 62, 4 (April 2005), 487–520. https://doi.org/10.1016/j.ijhcs.2004.12.002

[3] Leon Barnard, Ji Soo Yi, Julie A. Jacko, and Andrew Sears. 2007. Capturing the effects of context on human performance in mobile computing systems. *Personal and Ubiquitous Computing* 11, 2 (Jan. 2007), 81–96. https://doi.org/10.1007/s00779-006-0063-x

[4] Patrick Baudisch and Gerry Chu. 2009. Back-of-device interaction allows creating very small touch devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1923–1932.

[5] Joanna Bergstrom-Lehtovirta, Antti Oulasvirta, and Stephen Brewster. 2011. The effects of walking speed on target acquisition on a touchscreen interface. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*. 143–146.

[6] Meera M. Blattner, Denise A. Sumikawa, and Robert M. Greenberg. 1989. Earcons and Icons: Their Structure and Common Design Principles. *Human-Computer Interaction* 4, 1 (1989), 11–44.

[7] S. Brewster. 2002. Overcoming the Lack of Screen Space on Mobile Computers. *Personal and Ubiquitous Computing* 6, 3 (2002), 188–205.

[8] Stephen Brewster. 2002. Overcoming the lack of screen space on mobile computers. *Personal and Ubiquitous computing* 6, 3 (2002), 188–205.

[9] Stephen Brewster, Joanna Lumsden, Marek Bell, Malcolm Hall, and Stuart Tasker. 2003. Multimodal'eyes-free'interaction techniques for wearable devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 473–480.

[10] S. Brewster, P. Wright, and A.. Edwards. 1992. A Detailed Investigation into the Effectiveness of Earcons. In *International Conference on Auditory Display (Auditory display, sonification, audification and auditory interfaces, SFI Studies in the Sciences of Complexity, Vol. XVIII)*, Gregory Cramer (Ed.). Addison-Wesley, Santa Fe, 471–498.

[11] Stavros Garzonis, Simon Jones, Tim Jay, and Eamonn O'Neill. 2009. Auditory icon and earcon mobile service notifications: intuitiveness, learnability, memorability and preference. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1513–1522.

[12] W. W. Gaver. 1989. The SonicFinder: An Interface That Uses Auditory Icons. *Human-Computer Interaction* 4 (1989), 67–94.

[13] Mayank Goel, Leah Findlater, and Jacob Wobbrock. 2012. WalkType: using accelerometer data to accomodate situational impairments in mobile touch screen text entry. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Austin Texas USA, 2687–2696. https://doi.org/10.1145/2207676.2208662

[14] Mayank Goel, Jacob Wobbrock, and Shwetak Patel. 2012. GripSense: using built-in sensors to detect hand posture and pressure on commodity mobile phones. In *Proceedings of the 25th annual ACM symposium on User interface software and technology - UIST '12*. ACM Press, Cambridge, Massachusetts, USA, 545. https://doi.org/10.1145/2380116.2380184

[15] Jorge Goncalves, Zhanna Sarsenbayeva, Niels van Berkel, Chu Luo, Simo Hosio, Sirkka Risanen, Hannu Rintamäki, and Vassilis Kostakos. 2017. Tapping task performance on smartphones in cold temperature. *Interacting with Computers* 29, 3 (2017), 355–367.

[16] Michal Gordon and Cynthia Breazeal. 2015. Designing a virtual assistant for in-car child entertainment. In *Proceedings of the 14th International Conference on Interaction Design and Children*. 359–362.

[17] Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage publications Sage CA: Los Angeles, CA, 904–908.

[18] Morgan Harvey and Matthew Pointon. 2017. Perceptions of the effect of fragmented attention on mobile web search tasks. In *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval*. 293–296.

[19] Jibo He, Alex Chaparro, Bobby Nguyen, Rondell Burge, Joseph Crandall, Barbara Chaparro, Rui Ni, and Shi Cao. 2013. Texting while driving: Is speech-based texting less risky than handheld texting?. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 124–130.

[20] T. Hermann, A. Hunt, and John G. Neuhoff. 2011. *The sonification handbook*. Logos Verlag Berlin, GE.

[21] Matthew B Hoy. 2018. Alexa, Siri, Cortana, and more: an introduction to voice assistants. *Medical reference services quarterly* 37, 1 (2018), 81–88.

[22] Yoko Ishigami and Raymond M Klein. 2009. Is a hands-free phone safer than a handheld phone? *Journal of safety research* 40, 2 (2009), 157–164.

[23] Razan Jaber and Donald McMillan. 2020. Conversational user interfaces on mobile devices: Survey. In *Proceedings of the 2nd Conference on Conversational User Interfaces*. 1–11.

[24] Satu Jumisko-Pyykkö and Teija Vainio. 2010. Framing the context of use for mobile HCI. *International journal of mobile human computer interaction (IJMHCI)* 2, 4 (2010), 1–28.

[25] Marcel Adam Just, Timothy A Keller, and Jacquelyn Cynkar. 2008. A decrease in brain activation associated with driving when listening to someone speak. *Brain research* 1205 (2008), 70–80.

[26] Shaun K. Kane, Jacob O. Wobbrock, and Ian E. Smith. 2008. Getting off the treadmill: evaluating walking user interfaces for mobile devices in public spaces. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services - MobileHCI '08*. ACM Press, Amsterdam, The Netherlands, 109. https://doi.org/10.1145/1409240.1409253

[27] Amy K Karlson and Benjamin B Bederson. 2007. ThumbSpace: generalized one-handed input for touchscreen-based mobile devices. In *IFIP Conference on Human-Computer Interaction*. Springer, 324–338.

[28] Steinar Kristoffersen and Fredrik Ljungberg. [n. d.]. Mobile use of IT.

[29] Der-Song Lee, Kong-King Shieh, Shie-Chang Jeng, and I-Hsuan Shen. 2008. Effect of character size and lighting on legibility of electronic papers. *Displays* 29, 1 (2008), 10–17.

[30] Min Lin, Rich Goldman, Kathleen J. Price, Andrew Sears, and Julie Jacko. 2007. How do people tap when walking? An empirical investigation of nomadic data entry. *International Journal of Human-Computer Studies* 65, 9 (Sept. 2007), 759–769. https://doi.org/10.1016/j.ijhcs.2007.04.001

[31] Peter Liu, Fahad Zafar, and Aldo Badano. 2014. The effect of ambient illumination on handheld display image quality. *Journal of digital imaging* 27, 1 (2014), 12–18.

[32] Kent Lyons, Thad Starner, Daniel Plaisted, James Fusia, Amanda Lyons, Aaron Drew, and EW Looney. 2004. Twiddler typing: One-handed chording text entry for mobile phones. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 671–678.

[33] G. Marentakis and S. Brewster. 2005. Effects of reproduction equipment on interaction with a spatial audio interface. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems* (Portland, OR, USA). ACM, New York, NY, USA, 1625–1628. https://doi.org/10.1145/1056808.1056982

[34] Georgios N Marentakis and Stephen A Brewster. 2006. Effects of feedback, mobility and index of difficulty on deictic spatial audio target acquisition in the horizontal plane. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. 359–368.

[35] Joshua D McKeever, Maria T Schultheis, Vennila Padmanaban, and Allison Blasco. 2013. Driver performance while texting: even a little is too much. *Traffic injury prevention* 14, 2 (2013), 132–137.

[36] Martez E. Mott and Jacob O. Wobbrock. 2019. Cluster Touch: Improving Touch Accuracy on Smartphones for People with Motor and Situational Impairments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–14. https://doi.org/10.1145/3290605.3300257

[37] Daniel Munger, Bruce Mehler, Bryan Reimer, Jonathan Dobres, Anthony Pettinato, Brahmi Pugh, and Joseph F Coughlin. 2014. A simulation study examining smartphone destination entry while driving. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 1–5.

[38] Terhi Mustonen, Maria Olkkonen, and Jukka Hakkinen. 2004. Examining mobile phone text legibility while walking. In *CHI'04 extended abstracts on Human factors in computing systems*. 1243–1246.

[39] Judith Mwakalonge, Saidi Siuhi, and Jamario White. 2015. Distracted walking: Examining the extent to pedestrian safety problems. *Journal of traffic and transportation engineering (English edition)* 2, 5 (2015), 327–337.

[40] Alexander Ng, Stephen A Brewster, and John Williamson. 2013. The impact of encumbrance on mobile interactions. In *IFIP Conference on Human-Computer Interaction*. Springer, 92–109.

[41] Alexander Ng, Stephen A Brewster, and John H Williamson. 2014. Investigating the effects of encumbrance on one-and two-handed interactions with mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1981–1990.

[42] Alexander Ng, John Williamson, and Stephen Brewster. 2015. The Effects of Encumbrance and Mobility on Touch-Based Gesture Interactions for Mobile Phones. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 536–546. https://doi.org/10.1145/2785830.2785853

[43] Antti Oulasvirta. 2005. The fragmentation of attention in mobile interaction, and what to do with it. *interactions* 12, 6 (2005), 16–18.

[44] Antti Oulasvirta, Sakari Tamminen, Virpi Roto, and Jaana Kuorelahti. 2005. Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 919–928.

[45] Antti Pirhonen, Stephen Brewster, and Christopher Holguin. 2002. Gestural and audio metaphors as a means of control for mobile devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 291–298.

[46] Kathleen J. Price, Min Lin, Jinjuan Feng, Rich Goldman, Andrew Sears, and Julie A. Jacko. 2006. Motion does matter: an examination of speech-based text entry on the move. *Universal Access in the Information Society* 4, 3 (March 2006), 246–257. https://doi.org/10.1007/s10209-005-0006-8

[47] Ahmad Rahmati, Clayton Shepard, and Lin Zhong. 2009. NoShake: Content stabilization for shaking screens of mobile devices. In *2009 IEEE International Conference on Pervasive Computing and Communications*. IEEE, 1–6.

[48] T.V. Raman. 1997. *Auditory Interfaces: Towards the speaking computer*. Kluwer Academic Publishers.

[49] Zhanna Sarsenbayeva, Jorge Goncalves, Juan García, Simon Klakegg, Sirkka Rissanen, Hannu Rintamäki, Jari Hannu, and Vassilis Kostakos. 2016. Situational impairments to mobile interaction in cold environments. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. Association for Computing Machinery, New York, NY, USA, 85–96. https://doi.org/10.1145/2971648.2971734

[50] Zhanna Sarsenbayeva, Vassilis Kostakos, and Jorge Goncalves. 2019. Situationally-Induced Impairments and Disabilities Research. (2019), 5.

[51] Zhanna Sarsenbayeva, Niels van Berkel, Chu Luo, Vassilis Kostakos, and Jorge Goncalves. 2017. Challenges of situational impairments during interaction with mobile devices. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction*. ACM, Brisbane Queensland Australia, 477–481. https://doi.org/10.1145/3152771.3156161

[52] Zhanna Sarsenbayeva, Niels van Berkel, Aku Visuri, Sirkka Rissanen, Hannu Rintamaki, Vassilis Kostakos, and Jorge Goncalves. 2017. Sensing Cold-Induced Situational Impairments in Mobile Interaction Using Battery Temperature. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (Sept. 2017), 98:1–98:9. https://doi.org/10.1145/3130963

[53] Vibha Sazawal, Roy Want, and Gaetano Borriello. 2002. The unigesture approach one-handed text entry for small devices. In *International Conference on Mobile Human-Computer Interaction*. Springer, 256–270.

[54] Bastian Schildbach and Enrico Rukzio. 2010. Investigating Selection and Reading Performance on a Mobile Phone While Walking. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services* (Lisbon, Portugal) *(MobileHCI '10)*. Association for Computing Machinery, New York, NY, USA, 93–102. https://doi.org/10.1145/1851600.1851619

[55] Albrecht Schmidt, Kofi Asante Aidoo, Antti Takaluoma, Urpo Tuomela, Kristof Van Laerhoven, and Walter Van de Velde. 1999. Advanced Interaction in Context. In *Handheld and Ubiquitous Computing*, Hans-W. Gellersen (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 89–101.

[56] Andrew Sears and Mark Young. 2002. *Physical Disabilities and Computing Technologies: An Analysis of Impairments*. L. Erlbaum Associates Inc., USA, 482–503. https://doi.org/citation.cfm?id=772105

[57] Timothy Sohn, Kevin A Li, William G Griswold, and James D Hollan. 2008. A diary study of mobile information needs. In *Proceedings of the sigchi conference on human factors in computing systems*. 433–442.

[58] Garreth W. Tigwell, Zhanna Sarsenbayeva, Benjamin M. Gorman, David R. Flatla, Jorge Goncalves, Yeliz Yesilada, and Jacob O. Wobbrock. 2019. Addressing the Challenges of Situationally-Induced Impairments and Disabilities in Mobile Interaction. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–8. https://doi.org/10.1145/3290607.3299029

[59] Maciej Tomczak and Ewa Tomczak. 2014. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. (2014).

[60] Colton J Turner, Barbara S Chaparro, and Jibo He. 2021. Typing on a smartwatch while mobile: a comparison of input methods. *Human factors* 63, 6 (2021), 974–986.

[61] Kristin Vadas, Nirmal Patel, Kent Lyons, Thad Starner, and Julie Jacko. 2006. Reading On-the-Go: A Comparison of Audio and Hand-Held Displays. In *Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services* (Helsinki, Finland) *(MobileHCI '06)*. Association for Computing Machinery, New York, NY, USA, 219–226. https://doi.org/10.1145/1152215.1152262

[62] Bruce N Walker, Jeffrey Lindsay, Amanda Nance, Yoko Nakano, Dianne K Palladino, Tilman Dingler, and Myounghoon Jeon. 2013. Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus. *Human Factors* 55, 1 (2013), 157–182.

[63] Christopher D Wickens. 2002. Multiple resources and performance prediction. *Theoretical issues in ergonomics science* 3, 2 (2002), 159–177.

[64] Christopher D Wickens. 2008. Multiple resources and mental workload. *Human factors* 50, 3 (2008), 449–455.

[65] Christopher D Wickens, Diane L Sandry, and Michael Vidulich. 1983. Compatibility and resource competition between modalities of input, central processing, and output. *Human factors* 25, 2 (1983), 227–248.

[66] Jacob O. Wobbrock. 2019. Situationally-Induced Impairments and Disabilities. In *Web Accessibility: A Foundation for Research*, Yeliz Yesilada and Simon Harper (Eds.). Springer London, London, 59–92. https://doi.org/10.1007/978-1-4471-7440-0

[67] Jacob O Wobbrock, Shaun K Kane, Krzysztof Z Gajos, Susumu Harada, and Jon Froehlich. 2011. Ability-based design: Concept, principles and examples. *ACM Transactions on Accessible Computing (TACCESS)* 3, 3 (2011), 1–27.

[68] Tetsuo Yamabe and Kiyotaka Takahashi. 2007. Experiments in mobile user interface adaptation for walking users. In *The 2007 International Conference on Intelligent Pervasive Computing (IPC 2007)*. IEEE, 280–284.

[69] Shengdong Zhao, Pierre Dragicevic, Mark Chignell, Ravin Balakrishnan, and Patrick Baudisch. 2007. Earpod: eyes-free menu selection using touch input and reactive audio feedback. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 1395–1404.