Running with Technology: Evaluating the Impact of Interacting with Wearable Devices on Running Movement

MATTHIAS SEUTER, Münster UAS, Germany and University of Münster, Germany
MAX PFEIFFER, University of Münster, Germany
GERNOT BAUER, Münster UAS, Germany
KAREN ZENTGRAF, University of Münster, Germany
CHRISTIAN KRAY, University of Münster, Germany

The use of wearable devices during running has become commonplace. Although there is ongoing research on interaction techniques for use while running, the effects of the resulting interactions on the natural movement patterns have received little attention so far. While previous studies on pedestrians reported increased task load and reduced walking speed while interacting, running movement further restricts interaction and requires minimizing interferences, e.g. to avoid injuries and maximize comfort. In this paper, we aim to shed light on how interacting with wearable devices affects running movement. We present results from a motion-tracking study (N=12) evaluating changes in movement and task load when users interact with a smartphone, a smartwatch, or a pair of smartglasses while running. In our study, smartwatches required less effort than smartglasses when using swipe input, resulted in less interference with the running movement and were preferred overall. From our results, we infer a number of guidelines regarding interaction design targeting runners.

CCS Concepts: • **Human-centered computing → Laboratory experiments**; **Ubiquitous and mobile computing design and evaluation methods**; *User interface design*;

Additional Key Words and Phrases: wearables, interaction, running, movement, evaluation, motion capture

ACM Reference Format:
Matthias Seuter, Max Pfeiffer, Gernot Bauer, Karen Zentgraf, and Christian Kray. 2017. Running with Technology: Evaluating the Impact of Interacting with Wearable Devices on Running Movement. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 101 (September 2017), 17 pages.
DOI: http://doi.org/10.1145/3130966

The research reported here was conducted in the OpenLab at Münster University, which was funded by the federal state of Northrhine-Westfalia, the German Research Foundation (DFG) and Münster University. We are also grateful to Marc de Lussanet and Lena Slupinski for their explanations and their support in setting up the motion capture system.
Authors' addresses: M. Seuter and G. Bauer, Software Engineering Lab, Münster University of Applied Sciences, Dortmunder Straße 47, 48155 Münster, Germany; email: matthias.seuter@fh-muenster.de; M. Pfeiffer and C. Kray, Institute for Geoinformatics, University of Münster, Heisenbergstraße 2, 48149 Münster, Germany; K. Zentgraf, Institute of Sport and Exercise Sciences, University of Münster, Horstmarer Landweg 62 b, 48149 Münster, Germany .
Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2017 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.
2474-9567/2017/9-ART101 $15.00
DOI: http://doi.org/10.1145/3130966

1 INTRODUCTION

Sports such as running are being reshaped by wearable devices that provide useful services and information during training and competition. Due to the overall popularity of running, there are many services available supporting runners' performance, motivation, or focus: wearables can provide feedback on running technique [11, 41] or motivate runners, e.g. by simulating having to escape from a Zombie attack (cf. Zombie Run [43]). Another example is to facilitate running 'together' while being geographically separated by means of smartphones and headphones for creating awareness of the remote runner [24]. Even drones can serve as a means to support runners (and free them from having to wear a device while they are running) [25]. In the commercial realm, smart-/sportswatches and smartphones are widely available, and early adopters have started to use smartglasses.

While this trend towards more sophisticated sports apps and devices holds many promises, e.g. of better performance or more enjoyment, there are also potential downsides. One key issue relates to the fact that the interactions with such devices can disturb runners on different levels. Current research in HCI usually evaluates interactive tasks during movement by assessing the interaction itself. They often analyze task load and walking speed additionally. While both may provide insights into the interaction, they neglect a defining property of running, i.e. that a specific pattern of movements is repeated by a runner to perform the sports activity. This pattern incorporates coordinated movements of many parts of the body including, in particular, both legs and both arms. Since interactions with wearables typically require movements of body parts (e.g. to touch a button), these device interactions can interfere with the running movement. This, in turn, can affect enjoyment of the run and potentially affect performance and/or increase the risk of injury, similar to when texting while walking [35].

The goal of our work reported here is, therefore, to identify and quantify the impact of device interaction on the running activity. Our work further constitutes a key step towards designing interactive technology for use while running that minimizes interferences with the running movement. We report on a motion-capture study that compares a smartphone, smartwatch, and a pair of smartglasses. Participants had to briefly interact with each device while running. Our hypothesis was that influences from different devices vary. We analyzed changes in movement, self-reported task load for the interaction, and qualitative feedback from the participants. Our main contributions are (1) the identification and detailed quantification of disturbances of the running movement resulting from the interaction with three different device types; and (2) guidelines and implications for the design of interactive technology for use while running. Our method to measure interference with running movement and the results we obtained through it can also inform the design of future evaluation methods to assess the disturbance of movement patterns result from device interaction.

2 RELATED WORK

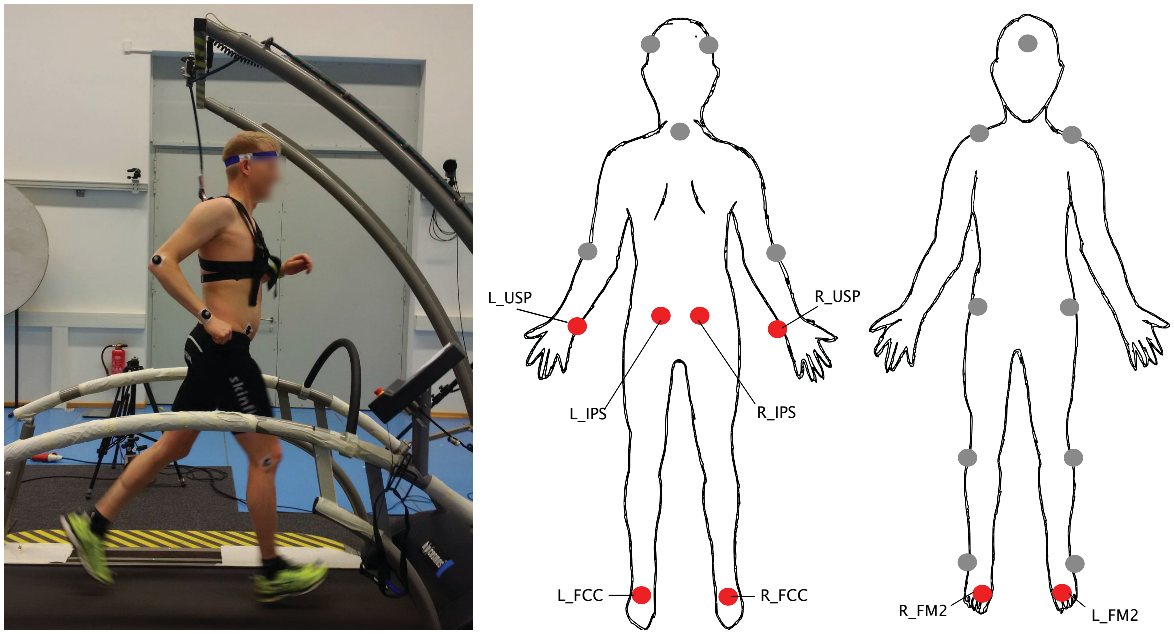
In addition to the opportunistic use of mobile apps while moving (e.g. texting while walking), sporting activities are a core application area for interactive technology intended for use while moving. Jensen et al. [15] propose a general design space for systems for use while running. Researchers have already explored different areas of this design space such as facilitating joint running over a distance [24], running with virtual characters [43], applications targeted at measuring and improving running performance [11, 41], or running with a drone [25]. Others have looked into visual feedback [30, 38] or video recording while running for storytelling [2]. Head mounted displays were used in [10, 42] for providing services for the runner that are always in sight. Still, active input may be required at certain moments, for example when switching between different services is needed. Such services usually require the user to interact at some point and potentially disturb the running activity.

This problem further aggravates by the fact that most interaction techniques involve movements that can interfere with physical activities such as running or walking. As a consequence of physical interference, Marshall and Tennent [22] noted that users often have to *stop to interact*, rather than being able to continue with their physical activity. They, thus, call for the development of interaction techniques that are feasible while the user is physically busy. Marshall et al. [21] further conceptualized the ordinal dimensions of the degree of the *relation between interaction and locomotion* and *inhibition of the interaction by locomotion.* In contrast, we aim to understand how interaction influences the running activity. In this sense, we extend their model and add another prospect by measuring the inhibition by the interaction on the movement.

Here, evaluations of influences on the running activity are sparse. Schneegas and Voit [36] provide one of the few recent evaluations of input techniques for the use while running. The authors focused on body-worn capacitive materials for sensing swipe gestures. Using camera recordings and questionnaires, they evaluated their textiles with users who ran a route outside. Smus and Kostakos [40] used foot gestures for changing a music track while running. They compared their gesture recognition to the mobile phone and a standard cable-button regarding timing and the user' preferences. These evaluations focused on the interaction and did not measure aspects of the running activity. Therefore, we next survey related work to identify promising measures and evaluation methods.

Walking and using mobile devices is a related research field since running as well as walking are forms of human gait. The history of evaluations of mobile interaction for use while walking is dominated by measured changes in walking speed and task load ratings using the NASA TLX (NASA task load index [12]) along with measures of input performance (e.g. timing). Pirohonen et al. were one of the first to use the PPWS (percentage of preferred walking speed). They compared different input modalities for a media player [31] with participants who walked along a predefined course and performed one input task on the media player per lap. They measured lap times to derive PPWS and used the NASA TLX to assess task load. Notable further examples focusing on PPWS and the NASA TLX include [4, 5, 27, 28, 34]. Other studies focused on the relation between the performance of using the device and walking, e.g. target timing and walking. In this area, Crossan et al. [7] measured gait oscillations using inertial sensors. They investigated how input behavior on a touch screen correlates with hand oscillations while walking. The authors found differences in accuracy and preference in when to touch. Crosson et al. [6, 8] also studied targeting via wrist rotation and head tilting when static and walking. Bergstrom-Lehtovirta et al. also studied how walking speed and targeting relate [3]. They found that targeting accuracy decreases as a function of walking speed. They further measured body and hand oscillation as an aspect of walking and found that target selection accuracy deteriorates in relation to both types of oscillation. Based on their findings, the authors emphasize the importance of considering the trade-off between interaction and movement. They also suggest the development of interaction techniques that improve this trade-off overall. We note that walking speed is a commonly used measure for reflecting on the effect of the interaction's task load on walking. Therefore, walking speed does not reflect changes in movement well.

Highly developed motion capture technologies are commonly used in other research areas to investigate changes in movement and gait. Modern motion capture techniques [14, 23] allow researchers to measure and infer many movement and force parameters. Research fields such as athletics or medical research are assessing gait and biomechanics by measuring, e.g. joint angles, body segment rotations, or stride behavior. Researchers recently adopted such measures for evaluating impacts of using devices on walking. Contributions to this discussion include [1, 18–20, 26, 32, 35]. Amongst other factors, gait stability for using mobile applications safely is frequently discussed, potentially preventing accidents and injuries in pedestrians. For example, medial-lateral foot deviation was used by Schabrun et al. [35] for reflecting on gait stability while walking. Often discussed are the origins of measured changes as they can depend on locomotion speed, which can change for many reasons. One issue with observed changes is that they potentially display small effects only. Here differences are often small and difficult to measure because gait is a long learned task in adult humans. For measuring changes in gait we also note that pedestrians who use devices while walking are commonly involved in longer tasks (e.g. texting or searching information about a location), compromising attention and time. In contrast to walking, interactions while running are short and compromises are undesirable. We, therefore, expect changes in the moment of the interaction only.



(a) Experimental setup

(b) Body landmarks

Fig. 1. (a) Participant running on the treadmill during experiment with safety belt and marker set attached; (b) Overview of our custom marker set. The named markers are used in our study.

Consequently, there is a need for evaluating wearable devices regarding their influence on movement and surveyed measures from general mobile computing and measures of gait.

3 MEASURING EFFECTS OF INTERACTING WITH WEARABLE DEVICES ON RUNNING

With the growing number of available wearable devices, we identified the need for evaluating their influence on the running activity when interacting with them. A method to compare different device types consequently addresses the following main requirements. First, the experimental task represents a complete interaction to facilitate exploring the new field. Second, the method must control other factors for comparing different devices. The main factors to be controlled are the task load incurred by the experimental task and the participants' constant physical effort.

There are many occasions for active interaction during running such as accepting calls, closing reminders, changing music, changing view, or changing running/monitoring parameters. For simulating the interaction task, upon hearing an audio signal, participants had to perform a swipe gesture on the device to request a symbol that is then shown on the device's display. A swipe gesture was available on all three devices (a tap gesture was not available on one of the used devices). The task further includes observing and remembering the shown symbols to ensure that the participants attend to the display.

We adopt objective measures from the literature that represent gait stability (cf. e.g. [35]) and also analyzed changes in the running movement in terms of *magnitude* and *duration* of deviations from the normal movement. Previous work has also used mental task load (cf. e.g. [31, 34]) for walking and interacting-studies. We use task load ratings regarding the remembering as a measure of the effect on the interaction incurred by using a device while running. A final quantitative measurement (cf. e.g. [31]) constitutes a set of open questions for assessing the users' demographics, opinions and feelings.

The basic kinematics of human gait are described by e.g. [29]. Here, to describe gait, parameters and phases are defined that depend on the timing of the foot striking the ground. We use optical motion capturing for tracking retro-reflective markers placed on the participants' skin [14] and the algorithm by Smith et al. [39] to detect foot strike points and for segmenting the data into steps. The foot strike detection requires markers on the feet (FCC and FM2) and hip (IPS). Comparing movements considering the interaction and the running requires markers on the hands and feet. Based on these considerations, we referred to the Helen-Hayes marker set [16], which was developed for clinical gait analysis and has been used in a large number of studies. We tested the marker positions for stable placement and practical use to derive our custom marker set. Figure 1(b) shows an overview of the 23 markers we kept (back and front). In this work, we make use of the markers at the Ulna-Styloid Process (L_USP, R_USP), Posterior Superior Iliac Spine (L_IPS, R_IPS), Metatarsus 2nd Head (L_FM2, R_FM2), and Calcaneus-Aspect of the Achilles Tendon (L_FCC, R_FCC). The treadmill running required to use shoes and thus to place the FCC, and FM2 on the shoes, at the heel and centered above the toes tips.

To keep the mental task comparable across devices, we need to consider mental rotation that differs depending on the device placement on the body. Assuming that rotations below 90° have no additional effect [13], we can use symmetric symbols to control for this aspect. Our experiment app randomly shows a circle ● or a square ■ when a participant performs a swipe gesture after the audio signal. The app automatically returns to a white screen after three seconds. We chose to use n-back remembering [17] as a secondary task to adjust the difficulty of the remembering task. The participants had to request and observe a new symbol s_i with the signal. They needed to remember the last n=3 symbols, and for each new symbol s_i had to announce whether it was the same symbol as observed three times before s_{i-3}, i.e. by saying *yes* if $s_i = s_{i-3}$ and *no* otherwise. For example, the series (●, ■, ■, ●, ●, ■, ●) yields the correct verbal responses (-, -, -, *yes, no, yes, yes*). In a pilot study, we found that n=3 kept the mental task load comparable and the participants described the task as being difficult while also mostly remembering correctly. The pilot study also confirmed that the two symbols could be distinguished well.

We addressed other potential issues using a lab based setting. It allows for ruling out any external factors such as obstacles, weather, other people or uneven ground and the treadmill enables the use of optical motion capture technology and to analyze kinematics for a longer series of steps. We counterbalanced the conditions to address any order effects. In addition, we address possible influences due to familiarity with treadmill running and the used devices by reserving familiarization time during warm up and before each condition. Also, while running at individual speeds (self-chosen, relaxed) and breaks between the conditions avoid fatigue in the participants during the second and third condition, heart rate measurement allows for monitoring the constant physical effort of the participants. We use off-the-shelf devices with standard functionality and typical properties in terms of size, weight, and shape. All selected devices could run our Android app to facilitate the experiment. Participants could freely decide where to wear the device and which hand to use for interacting to ensure maximum comfort and naturalness. This decision, in turn, required us to mirror left and right markers so that all participants (virtually) used their right hand for input on the device.

Our experiment described in the next section uses the developed approach to compare a smartwatch, smartphone, and a pair of smartglasses (see Figure 2(a)) as common representatives for wearable devices.

4 EXPERIMENT

Interacting with different device types can interfere with running in various ways. It can affect gait parameters and changes in the movement trajectories of body landmarks, self-reported task load, and the runners' opinions. We hypothesize that effects on movement influence the runners' preferences.

(a) Wearable devices (b) Jet input field

Fig. 2. (a) Devices used in the experiment: smartglasses, smartwatch, smartphone (from left to right); (b) the optical touch field of the smartglasses that was used to detect swipe gestures

4.1 Design

To study the interference that different devices introduce, we implemented a dual-task experiment based on the above described method. The primary task was running on a treadmill with low-level effort for the participant. The secondary task was remembering symbols shown on a wearable device (n-back remembering task). A smartphone, smartwatch, and a pair of smartglasses represented three different wearable device types. We followed a within-subject design with a counterbalanced (assigned randomly) order of the devices. We measured the participants' movement on the treadmill and recorded their interactions with the devices as well as their responses to the secondary task. In addition, we monitored their heart rate for safety reasons and to ensure they were properly warmed up before the start of the trials. Questionnaires were filled in on paper. The faculty ethics board approved the study.

4.2 Participants

We initially recruited 22 participants from around campus and from local sports clubs. We had to discard data from seven participants who experienced random input on the smartglasses. These false positives occur with moving the glasses' optical touch sensor through the system's infrared spotlights. We also discarded motion data of low quality from three participants. For these participants, all motion captures, for at least one condition, failed according to our removal criteria described below. We could thus use data from 12 participants for the analysis of the motion data, and from 15 participants for the analysis of the questionnaires. Table 1 summarizes the demographics. While all participants run regularly, their ambitions varied from frequently to not competing. Seven participants reported visual impairments, but all participants were able to use the devices without visual aids. They reported having no problems seeing the circle or square. Six of them reported near-sightedness and, therefore, no problems seeing the symbols on the devices; one participant reported astigmatism but also reported no problems seeing the symbols in the experiment. We compensated volunteers for their participation. They could choose between receiving money or experiment hours for their study program. In the first case, we paid them five Euros for each started half-hour of their time. In the second case, we attested the time taken to participate.

	Motion Data	Questionnaire Data	
Sex	6/6	8/7	*m/f*
Age in years	27.9 (7.6)	26.8 (7.2)	*mean (sd)*
Competing per year	3.6 (4.5)	3.1 (4.2)	*mean (sd)*
Running hours per week	2.7 (2.1)	2.7 (2.1)	*mean (sd)*
Self-chosen speed in km/h	7.8 (0.8)	7.7 (0.8)	*mean (sd)*
Used Smartwatch*	0/12	1/12	*Number of participants*
Used Smartphone*	8/12	11/15	*Number of participants*
Used Smartglasses*	0/12	0/15	*Number of participants*
Used Treadmill*	7/12	10/15	*Number of participants*

Table 1. Summary of the participants' demographics. Reported was prior experience with * during running. These differences were addressed by the familiarization steps in the procedure.

4.3 Apparatus and Material

Kinematic data was collected in the motion lab (OpenLab) of the University of Münster. using an optical motion tracking system from Qualisys. We used 12 Oqus cameras directed at the treadmill to capture the movement of the participants. We accounted for hidden space, reflections, ambient light and marker occlusions. The capturing rate was set to 400Hz and the system was calibrated according to the vendor-prescribed procedure. We used retro-reflective markers of one centimeter diameter. Markers were placed at the locations depicted in Figure 1(b) using double sided tape. We used elastic fixation tape between the skin and the double sided tape to prevent markers from falling off when in contact with sweat.

A neoprene band was used to secure the smartglasses (Recon Jet) around the back of the head. The smartwatch (Sony Smartwatch 3) was worn on the wrist without any further means to attach it. The smartphone (Motorola XT1032) was worn on the upper arm using a standard armband (see Figure 2(a)). While the smartwatch and the smartphone were placed on the arm participants preferred, the smartglasses did not allow for any adjustments regarding sidedness (see Figure 2(b)). A regularly maintained treadmill (pulsar 3p h/p/cosmos) was used in the experiment. Heart rates were recorded using a Garmin HRM-Run™ sensor and a Garmin 920xt watch.

Task load ratings were collected regarding the n-back remembering task. We used the German translation of NASA TLX [37] in the raw format [12], i.e. only the sub-scales without weighting were required. Participants also had to fill in a custom questionnaire, which included questions about their general fitness and health, their running habits and prior experiences with the devices we used. The questionnaire also asked participants to rank the three devices according to their personal preference for the use in the experiment and the use while running.

4.4 Procedure

The overall procedure of our experiment consisted of five phases that each participant went through: preparation, warm-up, base-line, main study, and debriefing. Each phase consisted of multiple steps detailed in the following. The procedure took between 80 and 120 minutes. Much time was devoted to breaks, forms, and preparation. The participants were active for only 40 to 50 minutes including the breaks between the conditions.

During *preparation*, we first welcomed participants and provided them with general information and the consent forms. We allocated 15 minutes for reading and completing the consent forms. The experimenter answered any questions participants had and verified they understood the information provided, including what data we collect and how we intended to use it. After they changed into their sports clothes and shoes, we placed the markers, connected the heart rate sensor as well as the treadmill's safety belt set and then verified the motion tracking was working properly. We monitored their heart rate during the experiment to ensure constant effort of the participants. During the *warm-up* phase, we asked the participants to slowly warm up and get used to running on the treadmill. They had as much warm up-time as they wanted. The display of the treadmill was covered so the participants could not refer to the displayed speed. Participants adjusted the speed via the faster/slower buttons on the treadmill until they found their preferred running speed. We transitioned to the next phase when their heart rate did not rise any further. The experimenter documented their normal heart rate and self-chosen speed at the end of the warm-up phase.

During the next phase, *base-line*, we captured the marker trajectories for the normal movement (without any device used). Depending on the participant's fitness level, we recorded seven to ten captures per condition. Each capture recorded 15 seconds tracking data of the placed markers. We started captures every 30 seconds. The next phase was the *main study*, and consisted of three trials during which participants ran on the treadmill (see Figure 1(a)) and interacted with one of the three devices (see Figure 2(a)). All trials had the same structure. At the beginning of each trial, participants were asked to put on one of the devices. We explained the experimental task, showed them an example and allowed them to try it out until they had mastered the task. Only then, the experimenter set the treadmill to the previously chosen speed. The task started when the normal heart rate was reached. The signal sound played automatically every 30 seconds. For each signal, capturing starting shortly before the signal and 15 seconds of running movement was captured. The participants completed the same number of interactions for each device. At the end of a trial, the treadmill stopped and participants filled in the NASA TLX while still being on the treadmill. After completing the questionnaire, we offered them to take a short drinking break before starting the next trial.

The final phase of *debriefing* began after all three trials of the main study phase were completed. We disconnected the heart rate monitor and the treadmill safety belt from the participants and removed the markers from their skin. They then could shower and/or change into fresh clothes before completing the final questionnaire. Once they had completed the questionnaire, we debriefed them about our research and answered any questions they had. We then thanked participants for taking part in the experiment and compensated them for their time.

		One way repeated measures ANOVA						**Paired t-tests**	
		DFn	*DFd*	*F*	*p*	η^2_G	p_{pw}	p_{pg}	p_{wg}
Displacement	max disp	2.00	22.00	9.35	0.00 **	0.27	0.02 *	0.00 **	1.00
	max xdisp	2.00	22.00	19.50	0.00 ***	0.35	0.00 ***	0.00 ***	1.00
	max ydisp	2.00	22.00	2.37	0.12	0.04	0.53	0.05	1.00
Outlier Counts	L_USP__R_FM2	1.25	13.74	36.92	0.00 ***	0.58	0.23	0.00 ***	0.00 ***
	R_USP__L_FM2	2.00	22.00	5.16	0.01 *	0.14	1.00	0.02 *	0.15
Max. Deviation (Right Steps)	L_USP__R_FM2	1.31	14.36	8.29	0.00 **	0.34	1.00	0.03 *	0.00 ***
	R_USP__L_FM2	2.00	22.00	17.16	0.00 ***	0.49	0.14	0.00 **	0.00 **
NASA TLX	Mental Demand	2.00	28.00	2.36	0.11	0.03	1.00	0.20	0.35
	Physical Demand	1.26	17.68	1.25	0.30	0.04	0.75	0.89	1.00
	Temporal Demand	1.40	19.59	0.96	0.40	0.01	1.00	0.82	0.29
	Performance	2.00	28.00	7.22	0.00 **	0.24	0.01 *	0.08	0.96
	Effort	2.00	28.00	4.50	0.02 *	0.05	0.12	1.00	0.02 *
	Frustration	2.00	28.00	2.77	0.08	0.07	1.00	0.10	0.63
	Overall	2.00	28.00	5.48	0.01 **	0.06	0.25	0.05 *	0.24
Task Performance	N-Back Remembering	2.00	28.00	2.06	0.15	0.07	1.00	0.55	0.25

Table 2. Output of the statistical tests for watch use w, phone use p, and glasses use g. The analyses revealed differences in maximum displacement (disp, xdisp), interference duration (outlier counts) and magnitude (maximum deviation from normal movement during right steps); self rated performance, effort and the overall task load. Analyses: One-way repeated measures ANOVA along with Mauchly's test for sphericity and Greenhouse-Geisser correction applied if required. Posthoc analyses: Paired t-tests; Bonferroni corrected p-values. Significance: *, **, *** indicating p<0.05, p<0.01, p<0.001.

5 ANALYSIS

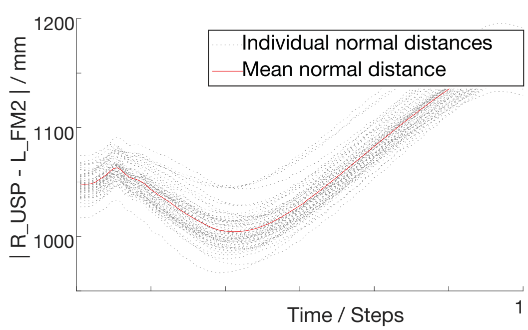
We observed data on three different levels. Our experiment allowed us to track movement data and to collect self reported task load after each condition. Participants also provided answers to a set of open questions after the experiment.

5.1 Data Preparation

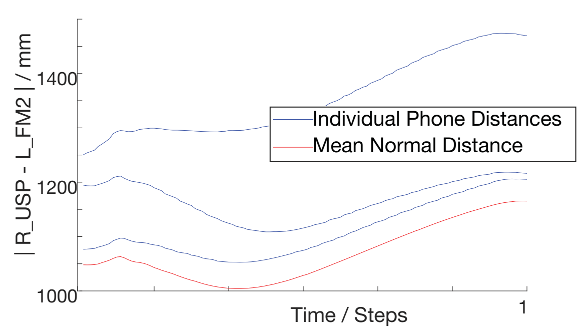
The raw motion data was preprocessed semi-automatically to facilitate the analyses. We used linear interpolation to fill gaps in identified trajectories smaller than ten frames. We reviewed gaps between 10 and 50 frames, and either used linear or polynomial interpolation to fit the trajectory. We removed the entire recording if there were any gaps larger than 50 frames. We also discarded the data for the three interactions with a device as they were only initiating the 3-back task. Finally, we manually removed captures that included movements that were clearly not related to the interaction or the running. For example, sometimes participants manually corrected the fitting of their trunks or other clothes.

5.2 Interference of Interaction with Running Movement

We calculated the displacement between the positions of two subsequent foot strikes of the same foot (FM2 markers). The displacement served as a measure of step behavior or *gait stability*. We analyzed the absolute displacement (disp), and the components (xdisp and ydisp). The xdisp measures displacement in the medial-lateral (left and right) direction (cf. [35]), and the ydisp measures displacement in back and forth-direction. For each condition, we subtracted the average maximum displacement of the normal behavior for comparing the conditions. We also calculated the step-cadence and approximated thigh and knee angles as the angles between the corresponding markers. However, their analyses did not show differences for the collected data.



(a) Normal



(b) Deviation from Normal

Fig. 3. For the right steps of participant p1021: (a) Shows the distance R_USP__L_FM2 of all normal steps and their resulting mean step, representing the participant's normal profile for this distance. (b) Shows three example steps when using a phone compared to the normal mean. The selection includes at least one deviating graph.

For analysing the deviation from the normal, we compare distances between markers. They provide relative data and reflect on the relation of the landmarks with each other. It makes sense to relate the left (right) hand with the right (left) foot because the arms cross-coat the legs in the human gait pattern and the corresponding limbs are normally in sync with each other. We thus calculated the distances between the markers R_USP__L_FM2 and R_FM2__L_USP. We next sampled 120 frames (minimum to fit our fastest participant) from each step linearly and plot the sampled steps per participant and condition. Figure 3 illustrates an example. It shows (a) the normal profile of participant p1021 represented by the distances between two body landmarks with all normal right steps and their mean as the reference; and (b) the normal mean together with three examples resulting from using a phone displaying visible deviations.

As a measure of *magnitude of changes*, we analyzed the peak deviations in the USP_FM2-distances. For this analysis, we subtracted the normal mean distance for all steps from the phone, watch and glasses conditions for each participant. We then calculated their maximum deviations for each condition. As a measure of *duration of changes*, we also analyzed the number outlying frames in the captures. We identified outlying frames in the USP_FM2-distances by comparing each step to the corresponding normal steps of the same participant. We tested a frame for being an outlier by first concatenating the distances to be tested with all corresponding normal distances and then testing the concatenated vector for outliers. The non-normal point in the vector was either marked outlying or not. We used Grubbs' [9] algorithm to identify outliers. Each of the two procedures result in one number per participant and condition for the statistical tests.

5.3 Self Reported Task Load for Remembering while Running

We used NASA TLX to collect subjective ratings regarding the remembering task. We did, therefore, not analyze differences between running with and without a device. We analyzed the individual subscales and their mean as the overall task load. All measures were then analyzed as described next. The description of the subscales also include the n-back remembering performance for comparison.

5.4 Statistical Tests

We used the mean for the statistical tests if multiple tracks led to more than one data point per participant. We ran a one way repeated measures ANOVA for each parameter. We applied Mauchly's test for sphericity and used Greenhouse Geisser corrections if necessary. We ran paired t-tests with Bonferroni correction as posthoc tests.

5.5 Feedback and Open Questions Addressing Preferences

The final questionnaire targeted preferences for *every day use* and *use in the experiment*. It also included a ranking task and a set of open questions. Two of the authors carefully coded the provided statements and identified clusters. We then named the resulting factors and counted their positive and negative occurrences.

6 RESULTS

We analyzed the data according to the steps outlined above. In the following, we describe our main findings. Table 2 reports the outcomes of the statistical tests and Figure 4 plots the found differences.

6.1 Interference of Interaction with Running Movement

The maximum displacement between two subsequent foot strikes (disp) was significantly larger for the phone condition compared to the watch ($p_{pw} = 0.02$), and the phone compared to the smartglasses ($p_{pg} = 0.00$). Also, sidewards-displacement (xdisp) was significantly larger for the phone compared to the smartglasses ($p_{pg} = 0.00$), and for the phone compared to the watch ($p_{pw} = 0.00$). Figure 4(a) illustrates these findings. Changes in the displacement in back and forth-direction (ydisp) were not observed.

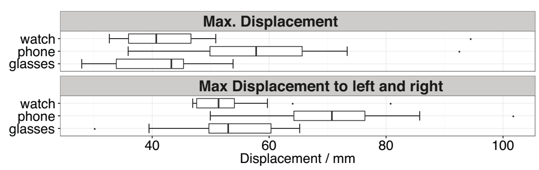
The analysis of the maximum distances between the right USP (at the hand doing the pointing movement) and the left FM2 revealed significant differences showing the influence of interacting with the smartglasses being higher than with the phone ($p_{pg} = 0.00$) and higher than with the watch ($p_{wg} = 0.00$). The analysis of the maximum distances between the left USP marker (the hand not doing the pointing movement) and the right FM2 also revealed significant differences. The impact of interacting with the phone on this distance was significantly higher than the with the smartglasses ($p_{pg} = 0.00$), and the impact of interacting with the watch was significantly higher than with the glasses ($p_{pw} = 0.03$). Figure 4(b) illustrates these findings. However, these findings apply for right foot steps only, we did not observe differences for left steps in these parameters.

Counting outliers in the marker distances revealed significant differences for the right USP to left FM2 distance (the distance between the hand doing the pointing movement and the opposite foot). Using the smartglasses while running resulted in significantly more outliers than using the smartphone ($p_{pg} = 0.02$). Counting outliers in marker distances also produced significant differences for the left USP to right FM2 distance (the distance between the hand not doing the pointing movement and the opposite foot). Here, the smartglasses condition showed significantly fewer outliers than the phone condition ($p_{pg} = 0.00$), and also significantly fewer outliers than the watch condition ($p_{wg} = 0.00$). Figure 4(c) illustrates these findings.

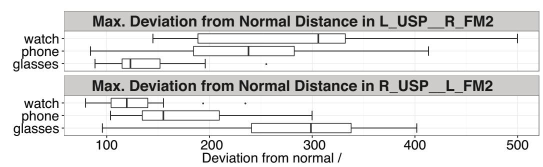
The described differences also show clear effect sizes and the corresponding mean differences are visible in the Figures 4(a), 4(b), and 4(c), which show the corresponding boxplots.

6.2 Self Reported Task Load for Remembering while Running

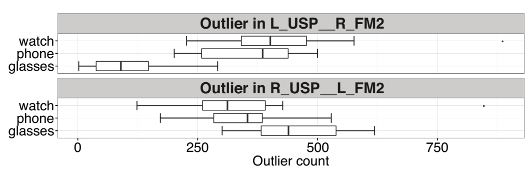
A one way repeated measures ANOVA and paired t-tests for the overall task load as reported in the NASA TLX questionnaire revealed that the smartglasses condition resulted in a significantly higher task load than the smartphone condition ($p_{pg} = 0.05$) overall. We observed two significant differences among the subscales. First, self-rated performance was higher when using the watch compared to using the phone ($p_{pw} = 0.01$). Second, significantly more effort was reported for using the smartglasses compared to using the watch ($p_{wg} = 0.02$). Figure 4(d) illustrates these findings. Our analysis did not show a difference for the remembering performance. Also, except for the self-rated performance, the analyses display small effect sizes only.



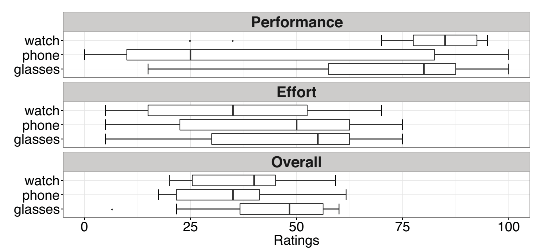
(a) Maximum displacement



(b) Maximum deviation from normal movement in distances between two markers



(c) Outlying frames compared to normal movement



(d) Self reported task load (NASA TLX)

Fig. 4. (a) Significant differences were observed for the maximum distances between two subsequent foot strike positions of the same foot. The phone shows more displacement than the watch and the glasses. (b) Significant differences were observed for the maximum deviation in the distances between the foot and the hand markers each. The pointing hand had to travel a significantly longer distance for the smartglasses, while the other hand to travel a significantly shorter distance. This difference applies to right steps only. (c) Significant differences were observed for outlier counts in distances defined by body landmarks that are more involved in the pointing movement. The plot shows the differences for the USP (hand) to opposite FM2 (foot) marker. (d) The NASA TLX ratings yield differences. The overall measure shows a significant difference between the glasses and the phone. The glasses also required significantly more effort compared to the watch.

6.3 Feedback and Open Questions Addressing Preferences

Answers from the final questionnaire indicated that participants preferred to use the watch in the experiment and during their running activities. We grouped responses into four categories. *Comfort* included placement and fit, e.g. comments such as *'... the smartphone bag did not feel stable.'*, size and weight of the carried device, and how practical the device was to wear. The *ease of use* category encompassed the management of the view field and the display observability, e.g. comments such as *'The targeting was difficult on the small input field of the smarglasses'*, or *'... the watch was easy to reach and to move into the view field.'*. Factors such as the input field size and simple, fast or practical usage were also included in this category. *Familiarity* included references to prior experiences, e.g. comments such as *'Familiar use. I always run with my watch'*, as well as general references to familiarity with device types, their use and interaction techniques. Finally, *interference* denoted mentions of additional movement required and effects on movement, for example, *'I needed to interdigitate my arms so that the interaction broke my rhythm'*. This category also included effects on the running rhythm, awkward device handling as well as visual, motor or mental distractions.

We also counted all comments per participant and category and classified them as either positive or negative. The counts are reported as (+/-) in the following. For the watch, comfort (5/0), ease of use (6/2), familiarity (3/0) and interference (4/1) were described most positively. The phone was frequently described negatively in comfort (1/3), ease of use (1/8), familiarity (1/3) and interference (1/7). The smartglasses received negative comments in comfort (0/3), ease of use (0/4), familiarity (0/4) and interference (0/9). The direct rankings reflected these findings. The watch received 11 votes for the first rank (phone: three, smartglasses:one) for use in the experiment and everyday-running use equally. The phone was ranked second (nine votes for use in the experiment and five votes for everyday-running use). The smartglasses were ranked last (10 and nine votes for third position).

7 DISCUSSION

Based on our findings and gathered experience during the study, we derive design implications for interactive run-technology and briefly discuss the limitations of our study and its results.

7.1 Interference of Device Interaction with Running Movement

Comparing the smartglasses to the phone or watch, we found that the pointing-hand, in relation to the opposite foot, deviates longer (number of outliers) and more (magnitude) from the normal movement. Conversely, the other hand deviated less for the glasses compared to the deviation for the phone or watch. Differences in the deviations from normal movement appeared one-sided only (for the pointing hand-side). Changes are either short or successfully countered by the participants. We also observed a small (i.e. about 2 cm) difference in the participants' maximum inter-step displacement for the phone, which increased compared to the other devices.

The change in the displacement could be due to the changed visual focus on the device. It is also possible that changes are balancing actions of the motor system, required to counter the arm and hand movement caused by the interaction. Another possible explanation is that adjusted body posture from the interaction with the device also influences interstep distances (see also [35]). While it is thus not clear how exactly this change develops in the motor system process, it is important to note that changes in gait can potentially impact running safety. Our finding is particularly interesting in this context as it shows that even very short interactions can affect the highly automated human gait movement.

While the measured effects are small, they align with the self reported task load as well as with device preferences of the participants. In our study, the smartwatch was rated as incurring the lowest workload, and since the task was the same for all three devices we tested, it makes sense to attribute this result to the device type and the runners' interaction with it. In line with this, the watch caused the least interference with the running movement and was preferred overall by the participants.

The answers to the open questions highlighted some of the potential reasons for the above-discussed changes in movement and the overall ratings. They included several statements about effects on the running movement, which we also measured. For example, while the glasses' display in sight was described as distracting, the watch was referred to as being more flexibly with respect to moving into the field of view. According to the participants, the glasses' (blind) targeting input was difficult, and the phone was described to inhibit normal movement too. This is reflected by our measurements, which also indicate that the phone and smartglasses have a stronger impact on the running movement than the watch.

7.2 Considerations for Design and Research

Measuring interference can support the design of wearable technology in addition to other aspects of wearability such as comfort. Our results provide initial evidence that it is desirable to *minimize disturbances of the running movement that are caused by interacting with technology during running*. Not only can this reduce potential risk of injury but data on self-reported task load, user preferences and answers to open questions all seem to support this statement. From a design perspective, it thus makes sense to include the degree and type of interference caused by a particular combination of wearable device and interaction mechanism in the design and evaluation process. Besides directly measuring disturbances as we did, it would be interesting to investigate more formal models that incorporate both, the movement for running *and* the device interaction.

There are several *options to minimize interferences caused by device interaction*. Our results highlight the type of device being used as one of them: the smartwatch, for example, caused less interference than the smartglasses. Clearly, the placement of the device on the body plays an important role here, e.g. reaching a phone strapped to the upper arm requires the opposing hand to travel further than to wrist-worn watch. However, few devices provide the flexibility to be placed at different locations on the body. In our study, we used the same type of interaction across all devices (swipe gesture), but other types of interaction hold the potential to reduce interferences too. For example, mid-air gestures might substantially reduce interferences if designed well. Voice interaction could also be an option but suffers from interference with breathing and people might also be reluctant to use this in public.

Despite the continuous visibility of the 'screen' of the smartglasses, they were the least preferred device option and caused more interferences than the phone and the watch. This finding indicates that *visual accessibility of information may be less important during running compared to the degree of interference with the movement* than intuition suggests. Though these are initial findings and subject to limitations discussed below, they can inform the design of future running support systems by helping to prioritize aspects. For example, when faced with the choice of placing a device in a location where it is easy to see but harder to reach, prioritizing reachability over visibility might be advisable. However, there are some boundary conditions to consider, in particular regarding continuous movement. In particular, if runners have to look away from the direction they are running in for too long, accidents might happen. To this end, the experiment did not show a decrease in the reported gait stability parameter with watch use. Participants also reported they were distracted by the display always in sight, which can lead to accidents equally likely.

Finally, we also gained some insights with respect to *how to measure disturbances of running movement*. We opted to use a very reliable but high-effort approach (i.e. motion-capture using a high-accuracy multi-camera system) in order to detect as many interferences as possible. Based on our experiences in designing and implementing our measurement method, we can note that a small number of markers (i.e. on the foot and hand for capturing gait and the introduced interaction-movement) was particularly useful for detecting disturbances. Furthermore, we consider the direct analysis of changes in movement (e.g. the deviation from normal movement) to be a good candidate to implement to capture with less accurate set-ups and less effort too. In contrast, the difficulties of measuring changes in gait currently require advanced motion capture methods for measuring small changes. The related work section also outlined this difficulty. Parameters that require less accuracy would allow for smaller marker sets or even markerless motion capturing. This would greatly reduce the set-up and effort for studies aiming at investigating interaction during running. Given the high effort and cost of our set-up, the development of low-cost measurement approaches (e.g. based on strategically placed IMUs) is a logical next step.

7.3 Limitations

While the number of participants was comparatively small, it proved to be sufficient to support our goal to explore movement interferences. The collected dataset includes a broad range of parameters that we could explore as opposed to collecting a bigger dataset for confirming a single dependency. Nevertheless, a larger cohort of participants could result in deeper insights. A second limitation relates to the use of a treadmill. While treadmill running does not differ considerably from overground running concerning kinematics [33], it still differs from running a route outside regarding many other factors (e.g. in terms of having to navigate the environment or adjusting to variations in the running surface). Similarly, the placed markers could have affected the naturalness of the running movement, though in principle they do not physically constrain movement (i.e. we removed those who restricted movement). No critical marker fall offs occurred in the analyzed participants because fixation tape addressed sweat and helper fixed potential problems during the breaks. The attached markers may have influenced participants. However, they reported they did not feel uncomfortable. Furthermore, to control running effort, heart rates were monitored. None of the participants showed signs of physical exhaustion or fatigue. Concentrating on the task was reported being difficult during the third condition. This could explain some of the displayed variation in the task load rankings. However, the conditions were counterbalanced to address this influence. To control interaction and to measure differences reliably, we asked participants to perform a very constrained and somewhat artificial task. It is possible that user behavior and preferences would differ running apps specifically targeted to support runners though the actual physical interaction could be the same as in our study. Nevertheless, these aspects limit the generalizability of our results and the inferred guidelines. Finally, it would have been desirable to measure forces in the runners' joints more directly as this would allow for inferring injury risks more precisely. However, for our study, we did not have access to an instrumented treadmill that was capable of measuring this aspect. Our experiment measured changes in movement apparent in the moment of the interaction (i.e. one or two steps of the runner). Therefore, the quantitative findings cannot display effects in relation to a whole running activity. Instead, the measures aim for reflecting on qualities of the interaction that are otherwise invisible or too short.

8 CONCLUSION

In this paper, we investigated how interacting with wearable devices affects running movement. We reported on a lab-based motion-capture study that had participants perform a swipe gesture on a smartphone, smartwatch, and a pair of smartglasses, while running on a treadmill. We found significant differences in movement changes between devices regarding gait stability as well as in magnitude and duration of interferences with the running movement. The analysis of self-reported task load also revealed significant differences between using devices. Overall, the smartwatch was not only preferred by participants explicitly but also scored favorably regarding interference with movement. Our results thus provide initial insights into how the interaction with different devices affect the running movement and can be used to reduce such interferences. Our design considerations provide guidance that designers can employ to design for minimal interference. Future interaction techniques can be tested with various devices as shown by our experiment. Our work thus contributes to avoiding changes in movement and, therefore, to increase comfort with using devices while running, and potentially also reduce the risk of injuries associated with reduced gait stability.

We compared different wearables for running using a whole task that simulates a whole interaction. From here, future research should dissect the interaction and focus on specific aspects such as targeting or observing the screen. An experiment can investigate the role of different positions on the body or different ways of and carrying the device such as holding the device in hand or attaching it using bands. Different input methods such as tap, swipe, or drawing a circle constitute another component to consider. Here, interaction performance should be measured and the timing aligned with the gait movement to further analyze the relation between movement and interaction. In the center of another possible future direction lies the question: What do changes in movement mean for the user? To address this question, we are considering further studies with other types of interactions (e.g. head or hand gestures) to minimize interference. Here, interferences may even serve as an input method when voluntarily generated by the users. Regarding the development of the evaluation method, investigating these issues with improved measures and further factors is one interesting direction for future research. This direction could employ an instrumented treadmill for estimating forces in the joints during the interaction while running. Another research direction would aim for simplifying and developing our method further. It would ultimately enable measuring interference with movement in the field with little effort.

REFERENCES

[1] Valentina Agostini, Francesco Lo Fermo, Giuseppe Massazza, and Marco Knaflitz. 2015. Does texting while walking really affect gait in young adults? *Journal of NeuroEngineering and Rehabilitation* 12, 1 (22 Sep 2015), 86.

[2] Albara Alohali, Kai Kunze, and Robert Earle. 2016. Run with Me: Designing Storytelling Tools for Runners. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp '16)*. ACM, New York, NY, USA, 5–8.

[3] 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 (MobileHCI '11)*. ACM, New York, NY, USA, 143–146.

[4] 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 (CHI '03)*. ACM, New York, NY, USA, 473–480.

[5] Enrico Costanza, Samuel A. Inverso, Rebecca Allen, and Pattie Maes. 2007. Intimate Interfaces in Action: Assessing the Usability and Subtlety of Emg-based Motionless Gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. ACM, New York, NY, USA, 819–828.

[6] Andrew Crossan, Mark McGill, Stephen Brewster, and Roderick Murray-Smith. 2009. Head Tilting for Interaction in Mobile Contexts. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '09)*. ACM, New York, NY, USA, 6:1–6:10.

[7] Andrew Crossan, Roderick Murray-Smith, Stephen Brewster, James Kelly, and Bojan Musizza. 2005. Gait Phase Effects in Mobile Interaction. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems (CHI EA '05)*. ACM, New York, NY, USA, 1312–1315.

[8] Andrew Crossan, John Williamson, Stephen Brewster, and Rod Murray-Smith. 2008. Wrist Rotation for Interaction in Mobile Contexts. In *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '08)*. ACM, New York, NY, USA, 435–438.

[9] Frank E. Grubbs. 1969. Procedures for Detecting Outlying Observations in Samples. *Technometrics* 11, 1 (Feb. 1969), 1–21.

[10] T. Hamada, M. Okada, and M. Kitazaki. 2017. Jogging with a virtual runner using a see-through HMD. In *2017 IEEE Virtual Reality (VR)*. 445–446.

[11] Tian Hao, Guoliang Xing, and Gang Zhou. 2015. RunBuddy: A Smartphone System for Running Rhythm Monitoring. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 133–144.

[12] 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, 904–908.

[13] Holger Heppe, Axel Kohler, Marie-Therese Fleddermann, and Karen Zentgraf. 2016. The Relationship between Expertise in Sports, Visuospatial, and Basic Cognitive Skills. *Movement Science and Sport Psychology* (2016), 904.

[14] Brian K. Higginson. 2009. Methods of Running Gait Analysis:. *Current Sports Medicine Reports* 8, 3 (2009), 136–141.

[15] Mads Møller Jensen and Florian 'Floyd' Mueller. 2014. Running with Technology: Where Are We Heading?. In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design (OzCHI '14)*. ACM, New York, NY, USA, 527–530.

[16] M. P. Kadaba, H. K. Ramakrishnan, and M. E. Wootten. 1990. Measurement of lower extremity kinematics during level walking. *Journal of Orthopaedic Research* 8, 3 (1990), 383–392.

[17] W. K. Kirchner. 1958. Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology* 55, 4 (April 1958), 352–358.

[18] Eric M. Lamberg and Lisa M. Muratori. 2012. Cell phones change the way we walk. *Gait & Posture* 35, 4 (2012), 688 – 690.
[19] Jongil Lim, Avelino Amado, Leo Sheehan, and Richard E. A. Van Emmerik. 2015. Dual task interference during walking: The effects of texting on situational awareness and gait stability. *Gait & Posture* 42, 4 (2015), 466–471.
[20] Jane R. Marone, Pooja B. Patel, Christopher P. Hurt, and Mark D. Grabiner. 2014. Frontal plane margin of stability is increased during texting while walking. *Gait & Posture* 40, 1 (2014), 243–246.
[21] Joe Marshall, Alexandru Dancu, and Florian "Floyd" Mueller. 2016. Interaction in Motion: Designing Truly Mobile Interaction. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*. ACM, New York, NY, USA, 215–228.
[22] Joe Marshall and Paul Tennent. 2013. Mobile Interaction Does Not Exist. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. ACM, Paris, France, 2069–2078.
[23] Thomas B. Moeslund, Adrian Hilton, and Volker Krüger. 2006. A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding* 104, 2 (2006), 90 – 126. Special Issue on Modeling People: Vision-based understanding of a person's shape, appearance, movement and behaviour.
[24] Florian Mueller, Frank Vetere, Martin R. Gibbs, Darren Edge, Stefan Agamanolis, and Jennifer G. Sheridan. 2010. Jogging over a Distance Between Europe and Australia. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology (UIST '10)*. ACM, New York, NY, USA, 189–198.
[25] Florian 'Floyd' Mueller and Matthew Muirhead. 2015. Jogging with a Quadcopter. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2023–2032.
[26] Jack L. Nasar and Derek Troyer. 2013. Pedestrian injuries due to mobile phone use in public places. *Accident Analysis & Prevention* 57 (Aug. 2013), 91–95.
[27] 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 (CHI '14)*. ACM, New York, NY, USA, 1981–1990.
[28] Alexander Ng, John H. Williamson, and Stephen A. Brewster. 2014. Comparing Evaluation Methods for Encumbrance and Walking on Interaction with Touchscreen Mobile Devices. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices; Services (MobileHCI '14)*. ACM, New York, NY, USA, 23–32.
[29] Tom F Novacheck. 1998. The biomechanics of running. *Gait & Posture* 7, 1 (1998), 77–95.
[30] Stina Nylander, Mattias Jacobsson, and Jakob Tholander. 2014. Runright: real-time visual and audio feedback on running. ACM Press, 583–586.
[31] 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 (CHI '02)*. ACM, New York, NY, USA, 291–298.
[32] Prudence Plummer, Sarah Apple, Colleen Dowd, and Eliza Keith. 2015. Texting and walking: Effect of environmental setting and task prioritization on dual-task interference in healthy young adults. *Gait & Posture* 41, 1 (Jan. 2015), 46–51.
[33] Patrick O. Riley, Jay Dicharry, Jason Franz, Ugo Della Croce, Robert P. Wilder, and D. Casey Kerrigan. 2008. A kinematics and kinetic comparison of overground and treadmill running. *Medicine and Science in Sports and Exercise* 40, 6 (June 2008), 1093–1100.
[34] Simon Robinson, Parisa Eslambolchilar, and Matt Jones. 2009. Evaluating Haptics for Information Discovery While Walking. In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology (BCS-HCI '09)*. British Computer Society, Swinton, UK, UK, 93–102.
[35] Siobhan M. Schabrun, Wolbert van den Hoorn, Alison Moorcroft, Cameron Greenland, and Paul W. Hodges. 2014. Texting and Walking: Strategies for Postural Control and Implications for Safety. *PLOS ONE* 9, 1 (Jan. 2014), e84312.
[36] Stefan Schneegass and Alexandra Voit. 2016. GestureSleeve: Using Touch Sensitive Fabrics for Gestural Input on the Forearm for Controlling Smartwatches. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers (ISWC '16)*. ACM, New York, NY, USA, 108–115.
[37] M. Sepehr. 1988. *NASA Task Load Index: deutsche Version*. Number 284 in Bericht. Institut für Fahrzeugtechnik, Technische Universität Berlin. Berlin, 1988.
[38] Matthias Seuter, Lucien Opitz, Gernot Bauer, and David Hochmann. 2016. Live-feedback from the IMUs: Animated 3D Visualization for Everyday-exercising. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp '16)*. ACM, New York, NY, USA, 904–907.
[39] Laura Smith, Stephen Preece, Duncan Mason, and Christopher Bramah. 2015. A comparison of kinematic algorithms to estimate gait events during overground running. *Gait & Posture* 41, 1 (2015), 39–43.
[40] Boris Smus and Vassilis Kostakos. 2010. Running Gestures: Hands-free Interaction During Physical Activity. In *Proceedings of the 12th ACM International Conference Adjunct Papers on Ubiquitous Computing - Adjunct (UbiComp '10 Adjunct)*. ACM, New York, NY, USA, 433–434.
[41] Christina Strohrmann, Julia Seiter, Yurima Llorca, and Gerhard Tröster. 2013. Can Smartphones Help with Running Technique?. In *Proceedings of the 4th International Conference on Ambient Systems, Networks and Technologies (ANT 2013), the 3rd International Conference on Sustainable Energy Information Technology (SEIT-2013), Halifax, Nova Scotia, Canada, June 25-28, 2013 (Procedia Computer Science)*,

Vol. 19. Elsevier, 902–907.

[42] Chek Tien Tan, Richard Byrne, Simon Lui, Weilong Liu, and Florian Mueller. 2015. JoggAR: A Mixed-modality AR Approach for Technology-augmented Jogging. In *SIGGRAPH Asia 2015 Mobile Graphics and Interactive Applications (SA '15)*. ACM, New York, NY, USA, 33:1–33:1. https://doi.org/10.1145/2818427.2818434

[43] Emma Witkowski. 2013. Running from Zombies. In *Proceedings of The 9th Australasian Conference on Interactive Entertainment: Matters of Life and Death (IE '13)*. ACM, New York, NY, USA, 1:1–1:8.

Received May 2017; revised July 2017; accepted July 2017.