

Leon Barnard · Ji Soo Yi  
Julie A. Jacko · Andrew Sears

## Capturing the effects of context on human performance in mobile computing systems

Received: 11 June 2004 / Accepted: 8 March 2005 / Published online: 14 April 2006  
© Springer-Verlag London Limited 2006

**Abstract** Many real world mobile device interactions occur in context-rich environments. However, the majority of empirical studies on mobile computing are conducted in static or idealized conditions, resulting in a deficit of understanding of how changes in context impact users' abilities to perform effectively. This paper attempts to address the disconnect between the actual use and the evaluation of mobile devices by varying contextual conditions and recording changes in behavior. A study was performed to investigate the specific effects of changes in motion, lighting, and task type on user performance and workload. The results indicate that common contextual variations can lead to dramatic changes in behavior and that interactions between contextual factors are also important to consider.

**Keywords** Context · Mobile computing · Human performance · Movement · Lighting · Walking

---

L. Barnard · J. S. Yi  
H. Milton Stewart School of Industrial and Systems Engineering,  
Georgia Institute of Technology, Atlanta, Georgia, USA  
E-mail: lbarnard@isye.gatech.edu  
E-mail: jyi@isye.gatech.edu

J. A. Jacko (✉)  
Wallace H. Coulter Department of Biomedical Engineering,  
Georgia Institute of Technology and  
Emory University School of Medicine,  
Atlanta, Georgia, USA  
E-mail: jacko@bme.gatech.edu  
Tel.: +1-404-3856751  
Fax: +1-404-3857364

A. Sears  
Interactive Systems Research Center UMBC,  
Baltimore, Maryland, USA  
E-mail: asears@umbc.edu

---

### 1 Introduction

#### 1.1 Motivation

Mobile computing is an empowering notion to many people. The value that can be added by enabling access to powerful electronic devices in a portable form has yet to be fully realized, but obviously has vast potential in terms of productivity and communication for business, personal, educational, and medical purposes, just to scratch the surface. However, there is currently a large gap between the vision of mobile computing that has been created and the existing state of mobile computing, due in part to the relative youth of the field, but also to the inherent challenge of designing devices that are intended to be mobile.

Compared to desktop computing, designing hardware and software for mobile computing presents a host of unique challenges, particularly because location, environment, connectivity, and other important factors are commonly unpredictable and dynamic. The strategies that have been demonstrated to be effective for desktop computing are only minimally useful for mobile computing. Clearly, different design and evaluation paradigms need to exist for mobile computing devices and environments. Brewster et al. [1] cite the inadequacy of the desktop metaphor for mobile computing for information presentation. This is merely a single example of the dissonance between effective desktop and mobile computing strategies. Johnson [2] goes on to say: "Generally speaking, HCI has developed a good understanding of how to design and evaluate forms of human computer interaction in 'fixed' contexts of use...This is not the situation of use for mobile computing" (p. 5). This brings up the issue of differences between desktop and mobile computing in terms of contexts of use. Kristoffersen and Ljungberg [3] make the point that for traditional desktop computing applications, the tasks take place within the computer, while for mobile computing, the tasks typically reside outside

of the computer (such as navigation, observation data recording). Thus, in many mobile computing interactions, there are multiple tasks taking place, often with the mobile task being secondary, which is why the context of use must be considered.

## 1.2 Understanding context and its effects

In this investigation, context is presumed to be a set of conditions or user states that influence the ways in which a human interacts with a mobile computing device. This is consistent with other user-centered definitions of context, as described in Oulasvirta [4], Beale and Lonsdale [5], and Dourish [6], among others. These definitions are distinct from other, more technology-centered definitions and investigations, such as those found in studies focused on contextual architecture and components such as location, time, and nearby objects (i.e., [7–11]).

A good model that considers multiple components of context and adequately addresses aspects relevant to the user was presented in Sears et al. [12], where previous models were combined and organized into a three-dimensional context space, as presented in Fig. 1.

Even though user-centered context is understood to be important in the design and evaluation of mobile devices, and some appliance design approaches even acknowledge that contextual factors, such as ambience and attention, are crucial in device and task design, a surprising number of studies on mobile computing ignore the context of use relevant to the user. In a summary of 102 mobile HCI research papers from 2000 to 2002, Kjeldskov and Graham [13] found “a clear bias towards environment independent and artificial setting research...at the expense of natural setting research focusing on real use...” (p. 326). These authors further state, referring to the same papers: “understanding and learning from the design and real use of systems is less prioritized, limiting the generation of a cumulative body of knowledge on mobile human–computer interaction.” (p. 326). The need for real use studies is supported by Brewster [14], who, after experimenting with mobile device evaluation in a somewhat realistic situation, noted “a more realistic environment can significantly change the interaction and this must be taken into account when designing and testing mobile devices” (p. 202). He further urges other researchers to employ

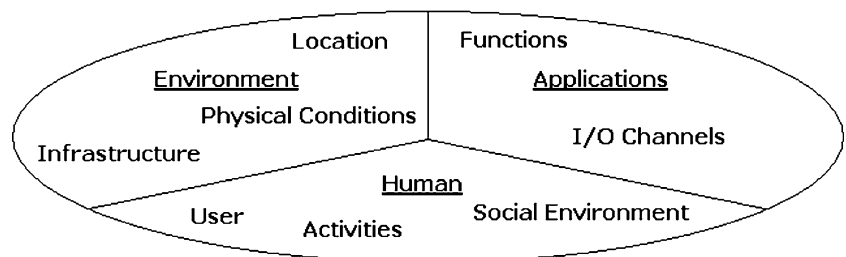
more appropriate evaluation strategies, while Johnson [2] states a need for new evaluation methods that are specific to mobile computing and specifies the demands of evaluating mobile systems as one of his four problems of HCI for mobile systems.

There is commonly a tradeoff between mobility and usability, in particular because users’ abilities are often hindered by their environment or situation. This can be viewed as the crux of the challenge of user-centered context-awareness. For example, increasing the text size may aid readability on a mobile device, but it also limits the amount of information that can be presented on a single screen and requires more scrolling to view the same amount of information. If this tradeoff did not exist, devices would remain in their optimal state at all times. But the optimal state for any mobile device is variable because of the wide variety of situations it is commonly used in, and is therefore context dependent.

The effect that environmental and contextual changes can have on mobile device users can be likened to the effects that physical or cognitive impairments can have on users with disabilities. For example, a blind person will typically have difficulty with an interface that was designed for use by a sighted person, like a touch screen. In the same way, a person entering text on a personal digital assistant (PDA) while walking will have difficulty doing so if the device has not been designed with consideration of a mobile user and device. The concept of situationally-induced impairments and disabilities (SIID) was introduced by Sears et al. [12] to describe some of the side-effects of working with a device in a situation that may impose constraints on the user’s ability to effectively accomplish their goals. The added dimension of variable conditions of use when using a mobile device means that the user may face unpredictable, and often less-than-ideal circumstances of use.

In order to sense and record relevant contextual factors, several investigations have been conducted to enable devices to sense characteristics of mobility, such as motion and changes in environmental conditions. Hinckley et al. [15] added proximity, touch, and tilt sensors to a commercially available PDA in order to allow the device to record important contextual information. Schmidt et al. [16] incorporated orientation and light sensors into a PDA device and specified additional sensors that could be used to retrieve information about conditions of the physical environment, such as acceleration, sound, and temperature. Using sensors to

**Fig. 1** Three-dimensional context space [from 12]



gather pertinent environmental information has been the focus of numerous recent research efforts (i.e., [17–23]).

Sensors can gather information relevant to context, such as location, acceleration, lighting levels, orientation, etc. However, being able to measure and/or record contextual factors that are relevant to the user is only the beginning. With sensors, contextual information can be collected, but the critical question is what to do with that information. The domain of context-awareness is nearing a state where it is faced with an abundance of potentially relevant available data, but a deficit of knowledge of how to use it. Designers may assume that these contextual factors are important, and even intuitively design with them in mind, but what is missing is an understanding of how changes in context affect the user. In most cases, a connection has not been made between the collected data and user behavior and performance. Bellotti and Edwards [24] provide an elegant anecdote as evidence: “a context aware application can measure temperature, but it cannot tell when a room is too hot and needs to be cooled” (p. 197). Information about how users react to changes in context and perform in context-rich environments is crucial to informing strategies for mobile device design.

### 1.3 Mobility as context

One of the most important aspects of context in mobile computing is mobility itself. It is variable, complex, and is highly pertinent to mobile computing. Kristoffersen and Ljungberg [25] define three types of mobility: traveling, visiting, and wandering. All three types frequently occur in novel environments. Wandering and traveling are similar, yet differentiated primarily by scale; wandering is conceived of as movement within a place whereas traveling is defined as movement between places. Visiting is the act of being in one place for a limited amount of time and then moving on to another place. With these three types of mobility only visiting implies a truly static position for any reasonable length of time. Traveling and wandering are predominantly spatially dynamic states, requiring applications that afford using while moving. Mobile computing can occur in any of these states of mobility, therefore it is essential to consider how interactions are performed in the traveling and wandering states, in particular, as they are the most challenging for mobile device designers. Currently, mobile computing devices and applications seem to be predominantly designed around use in visiting scenarios. A recent investigation cited movement as the most important contextual factor relevant to the user [26].

While it is intuitive that changes in mobility context affect a user’s interaction with a mobile device, very little is understood about how that interaction is affected. Even 6 years after Johnson, and others before him, raised the issue; it is surprising that most empirical studies of mobile devices, even those that discuss context explicitly, typically do not evaluate their designs

in mobile conditions. As a result of this, Dunlop and Brewster [27] have cited designing for mobility as the number one challenge to mobile device HCI designers. Yet despite, or perhaps because of, this, most mobile products are evaluated with users in static, highly controlled environments.

In addition to mobility, many other factors are relevant to a user who is interacting with a mobile device. In particular, the combination of mobility and other contextual factors is of interest, as multiple limiting factors are commonly present when the user is moving, such as excess noise, inadequate lighting, stress, inclement weather conditions, as well as changing tasks. One of the most common changing environmental conditions is lighting level. Changes in lighting are frequently encountered by mobile users, especially as they move from indoors to outdoors, but also as they move from room to room within a building or sun to shade while outside. Because information provided by mobile computing devices is almost exclusively visual, any condition that interferes with the visual salience of information displayed is important to examine. This factor is also listed as an important contextual identifier by Bristow et al. [26].

A few recent studies have looked at the context of mobility explicitly, examining either how motion affects the evaluation of mobile computing devices [28, 29] or how motion affects performance [30]. Additionally, Pascoe et al. [31] considered device requirements for mobile field workers.

The study described in this paper attempts to build upon this previous work as well as contribute a new level of rigor to the investigation of behavior in contextually rich environments, enabling deeper discovery of the specific effects of context on mobile device users. Ideally, this investigation will serve to show the benefits that can be obtained by investigating mobile devices in realistic contexts and convince other researchers to consider more realistic contexts during design and evaluation.

---

## 2 Methodology

### 2.1 Study objectives

In this study, three specific contextual factors (task type, motion, and lighting level) were manipulated in order to determine their relative effects on performance of mobile device users. While there has been abundant discussion of strategies for adapting mobile devices to changes in context, the degree to which changes in context impact a user’s ability to perform effectively are relatively unknown. Therefore, a clearer understanding of the effects of some of these changes in context on the user can help designers of context-aware tools better focus their efforts, and prioritize their context-sensing projects. The contextual factors studied here are intended to be representative of a subset of particularly relevant aspects of context, but are by no means exhaustive. The goal is to

establish a foundation by which the effects of context can begin to be more clearly understood.

## 2.2 Participants

One hundred twenty-six participants were asked to perform a set of tasks on a mobile device while sitting, walking on a treadmill, or free walking along a path around a room (the demographic information of the participants and the procedure of the experiment will be described below). Data from a subset of the participants, those who performed the tasks while sitting (“sitting group”) or while walking around the room (“walking group”), are examined in this paper. The treadmill condition is less relevant to this discussion, which centers on comparisons between absence and presence of mobility as a contextual factor, and was thus excluded from the analyses presented in this paper. The results from the treadmill condition have been analyzed separately and published elsewhere [32].

The participants considered in the present study ( $N=80$ ) volunteered over the course of one semester from an undergraduate Industrial Engineering course at Georgia Institute of Technology. The participants were primarily juniors and seniors, with a mean age of 21.8. Basic demographic characteristics were recorded for each participant. These demographic factors are summarized in Table 1. For age (in years) and PDA familiarity score (which was the sum of participants’ average frequency of use of common PDA applications and their overall PDA comfort level, on a scale from 0 to 8, where 0 represented no previous exposure at all and 8 indicated a very high level of expertise and familiarity), the mean value, along with the standard deviation in parentheses, are shown. Frequency counts are shown for the remaining factors.

As can be seen from Table 1, the majority of participants were not frequent PDA users, indicated by a high number of participants who had never owned a PDA and the relatively low average PDA familiarity score. The experimental tasks were performed on a PDA, however the tasks were designed such that prior experience with handheld devices was not required, as training was provided for each task and minimal input was needed to accomplish the goals of the tasks.

Each participant was randomly assigned to one of two groups: those who would be performing the tasks on a PDA while sitting, and those who would be performing the same tasks on a PDA while walking around a path

with obstacles within an observation room. The demographics for each group are presented in Table 2, along with the results of statistical tests comparing the groups in each category. Analysis of variance (ANOVA) was used to assess differences in age, while data not meeting the assumptions required for ANOVA were analyzed using either the Pearson chi-square test (when all categories contained five or more participants, a requirement for the Pearson chi-square test) or Kruskal Wallis chi-square test (when one or more categories contained less than five participants). Instances where the Kruskal Wallis chi-square test was used are designated with a superscript a.

The two groups were found to be statistically similar in all demographic characteristics except for native language, where a larger percentage of non-native speakers appeared in the Sitting group. The manner in which the difference in the number of non-native speakers between the two groups was accounted for will be discussed in Sect. 3.

## 2.3 Experimental tasks and conditions

Each participant performed two tasks in each of the two lighting conditions. Participants in the sitting group performed the tasks while sitting at a table, while participants in the walking group performed the tasks while walking along a path that contained obstacles.

For clarity, explanation of specific terms will be provided here. For the duration of the paper, the following words will match the specific meanings given below:

*Task* There were two tasks used in this study: (1) Reading Comprehension, and (2) Word Search. These were each separate activities with separate goals and instructions.

*Scenario* Within each task, participants performed two scenarios; a scenario was a combination of a task and lighting condition. Overall, there were four scenarios used: (1) Reading Comprehension + High-Light; (2) Reading Comprehension + Low-Light; (3) Word Search + High-Light; and (4) Word Search + Low-Light. Each scenario consisted of ten trials.

*Trial* A trial was defined to be a recorded action (such as answering a question) within a scenario. Each action was performed multiple times in order to get more

**Table 1** Demographics summary

$N$	Age	Gender	Native language	Dominant hand	Computer use frequency	PDA owner?	PDA familiarity score	Regularly read while walking?
80	21.8 (1.37) <sup>a</sup>	Male = 42 Female = 38	English = 66 Other = 14	Right = 76 Left = 4	> 1/day = 77 ≈ 1/day = 3 < 1/day = 0	Current = 13 Previous = 12 Never = 55	1.78 (1.69) <sup>a</sup>	Yes = 29 No = 51

<sup>a</sup>Average (SD)



**Table 2** Demographic comparisons between groups

Group	Sitting ( $N=42$ )	Walking ( $N=38$ )	Test statistic	$p$
Age	21.9 (1.45) <sup>b</sup>	21.7 (1.29) <sup>b</sup>	$F=0.405$	0.526
Gender	Male = 25; Female = 17	Male = 17; Female = 21	$\chi^2=1.74$	0.186
Native language	English = 31; Other = 11	English = 35; Other = 3	$\chi^2=4.56^a$	0.033*
Dominant hand	Right = 39; Left = 3	Right = 37; Left = 1	$\chi^2=0.84^a$	0.358
Computer use frequency	> 1/day = 41; $\approx$ 1/day = 1; < 1/day = 0	> 1/day = 36; $\approx$ 1/day = 2; < 1/day = 0	$\chi^2=0.45^a$	0.501
PDA owner?	Current = 8; Previous = 4; Never = 30	Current = 5; Previous = 8; Never = 25	$\chi^2=2.286$	0.319
PDA familiarity score	1.74 (1.69) <sup>b</sup>	1.82 (1.70) <sup>b</sup>	$\chi^2=0.02^a$	0.869
Regularly read while walking?	Yes = 15; No = 27	Yes = 14; No = 24	$\chi^2=0.01$	0.917

An asterisk (“\*\*”) denotes statistically significant at level 0.05

<sup>a</sup>Using Kruskal Wallis chi-square test

<sup>b</sup>Average (SD)

accurate measures of performance, as is common in empirical studies.

### 2.3.1 Lighting level

The light level in the observation room was used as a within-subjects variable with two levels. The room that was used for both sitting and walking conditions contained nine sets of overhead fluorescent lights, each with three bulbs. In the High-Light condition all 27 bulbs were illuminated, resulting in an intensity of approximately 260 lux; in the Low-Light condition only the middle bulb for each of the nine sets of lights was turned on, reducing the lighting to an average of 85 lux. For all tasks, the PDA backlight was turned off. The order in which each lighting condition occurred was randomized for each participant.

### 2.3.2 Task 1: Reading Comprehension

In order to assess a mobile device user’s ability to process information at a relatively deep level, a Reading Comprehension task was assigned to each participant. Given the widespread availability of eBooks [33] and other text document viewers for PDAs, reading comprehension was presumed to be of interest in the domain of mobile computing. Additionally, many other PDA tasks involve deep processing of information, such as the specific details of an appointment or meeting. The task involved reading paragraphs composed of fictional stories three to five sentences long and answering two multiple choice questions for each paragraph. The questions were taken from a book of standardized reading comprehension questions [34]. The tasks of reading and answering the multiple-choice questions were both carried out using a PDA. A Palm m505 PDA was used throughout the study for all participants. Participants read through five reading passages, each followed by two multiple choice questions in each of the two lighting conditions, for a total of ten passages of text and 20 questions. The same ten passages and 20 questions were used for all participants, but the order in

which they were presented was randomized. Screen shots of one of the text passages and one of the multiple choice screens are shown in Figs. 2 and 3.

Some scrolling was required for most of the reading passages and some of the multiple-choice questions, which could be done using either the up and down physical buttons on the device or by tapping small arrows on the screen with the stylus. After participants finished reading a text passage they pressed a button at the bottom of the screen labeled “Done”, which took them to the first of two questions about the passage they had just read. Participants were not allowed to go back to the passage once they had pressed the “Done” button and were therefore instructed not to move onto the questions until they felt they had sufficient understanding of the content of the passages. Once on a question screen, participants would see the question followed by four radio buttons next to four answer choices. A “Submit” button was at the very bottom of the screen and was used to submit the participant’s choice once they had selected an answer from the list. On the first multiple choice screen the “Submit” button would take the participant to a screen with the next question and answer choices, which would then take the participant directly to the next passage of text to read or to a screen which read “Task Finished” if it was the last question in the task.

### 2.3.3 Task 2: Word Search

Because many PDA tasks are shallower in nature than reading comprehension, such as looking up the phone number of a contact, or time of an appointment, a second task was used in the study in order to cover a broader spectrum of representative mobile device tasks. Again, the task was administered twice to each participant, once with all of the lights turned on and once with two-thirds of them turned off. Upon initiation of the task, participants were presented with a single screen of text chosen from a recent news article with a word that had been randomly selected from that text in larger letters at the top of the screen. The objective was to locate the word at the top of the screen (the “target”

word) within the text. For experimental soundness, the target word could only be a word of three letters or more and only appear once in the text. Once the participant located the target word they were instructed to tap anywhere on the line which contained the word with the



Fig. 2 Text passage screen shot

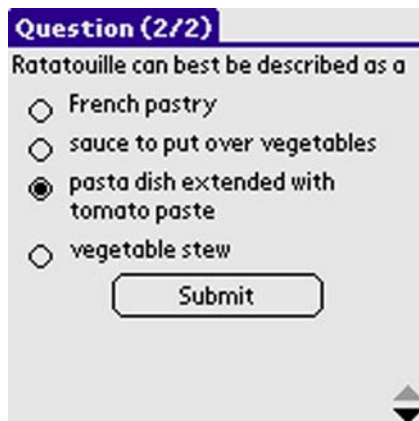


Fig. 3 Answer choice screen shot



Fig. 4 Word Search screen shot

stylus. This was done to reduce the stylus precision required by the user, equalizing differences in motor skills or PDA experience between participants sufficiently. Immediately upon touching the PDA screen, a new passage of text with a new target word would appear. This occurred whether the selection was correct or not. Participants performed this activity ten times in each lighting condition, for a total of 20 trials per participant. No participant saw the same text passage more than once. A sample screen shot for this task is shown in Fig. 4.

### 2.3.4 Condition 1: sitting

Participants assigned to the sitting group performed both tasks in both lighting conditions while sitting at a table in the observation room. They were provided with no specific instructions as to how to sit or whether or not to touch the table, only that they could not perform the tasks with the PDA resting flat on the table.

### 2.3.5 Condition 2: walking

Participants assigned to the walking group performed all tasks while walking around a 1-ft wide path that had been taped to a carpeted floor. The path was a loop that wound around tables and chairs in the room, such that users could make multiple laps during a single task scenario. The initial direction that the participants walked along the path was randomly chosen, and then alternated for the remaining three task scenarios of the experiment. The heavy black lines in Fig. 5 indicate the path that participants followed. The room was approximately 30 ft wide by 30 ft long and is shown to scale in

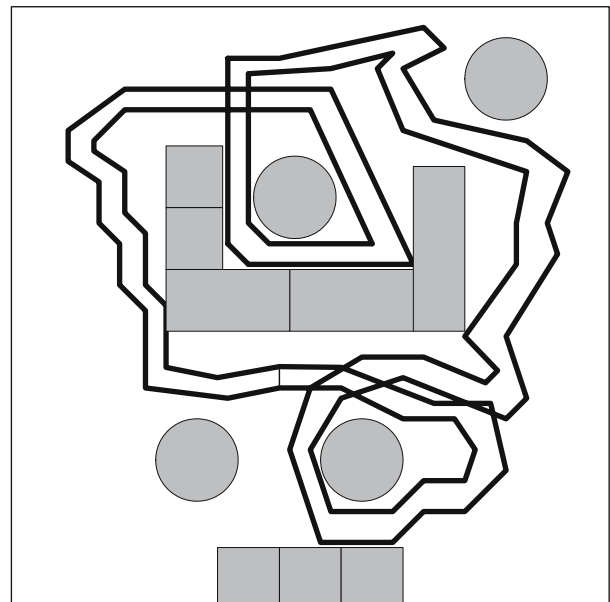


Fig. 5 The path followed by participants in the walking condition



Fig. 6 Demonstration of the walking scenario

the figure (except for the width of the tape which has been exaggerated for clarity).

Participants were instructed to keep both feet within the tape on either side of the path and informed that the number of times that they stepped on the tape would be recorded by the experimenter during the task. The number of full and partial laps that participants completed during each task scenario was also recorded by the experimenter, which was converted to distance (in feet) afterwards. There was no restriction on walking speed placed on the participants, only that they needed to keep moving. The walking condition is demonstrated in Fig. 6.

In summary, each participant was assigned to perform two tasks twice while either sitting or walking. The task they performed first (Reading Comprehension or Word Search) was selected at random for each participant. Once the task order was chosen, the participant performed the chosen task twice in each of two randomly assigned lighting condition scenarios. These two task scenarios were then followed by the other task in each of two lighting conditions.

### 2.3.6 Experimental apparatus

In addition to the Palm m505 PDA, several other tools were used in this study. In order to better understand

the differences in motion experienced by participants in each group, a triaxial accelerometer [35] was attached to the back of the PDA throughout the experiment. Accelerometer data ( $X$ ,  $Y$ ,  $Z$ , and “Net”, the vector sum of  $X$ ,  $Y$ , and  $Z$ ) were recorded after each task scenario, resulting in four separate records of acceleration for each participant. Additionally, a laptop computer was used to administer background and post-task surveys to the participants in addition to the NASA-TLX subjective workload assessment [36], which was administered after each of the four task scenarios.

## 2.4 Procedure

Before the experiment began, participants were given an introduction to the NASA-TLX workload assessment that was to be used in the study and given an opportunity to ask any questions about the meanings of the terms that were used. If the participant had been assigned to the walking condition, the next step was to determine a representative walking speed of that participant by having them walk two laps (one lap in each direction) around the path in the observation room. This was done in order to familiarize the participant with the path, as well as to establish a baseline walking speed to assess how much of an effect performing the task on a PDA had on their walking speed. Apart from this step, the procedures for participants in the sitting and walking groups were nearly identical.

Participants were given a verbal description of the first task they would be performing (either Reading Comprehension or Word Search), accompanied by text instructions on the PDA, and then given a chance to perform practice trials. In the case of reading comprehension, this consisted of one passage of text, followed by one multiple-choice question. For Word Search, the practice trials consisted of five instances of the task using text passages that were similar to the ones that would be used during the recorded trials. Practice trials were only given before the first scenario within each task. In the walking condition, participants performed the practice trials while walking around the taped path. Once participants verbally stated that they were comfortable with the task, the lighting level was adjusted to the scenarios at hand, and participants were instructed to begin the recorded trials.

Upon completion of the trials, participants filled out the NASA-TLX workload assessment. The lighting level was then adjusted to High-Light if the first scenario had been Low-Light or vice-versa. Participants then began the next set of trials for the same task. The NASA-TLX was then administered again, which completed the first task. Participants were then introduced to the next task and given practice trials before beginning. The procedure for the second task

**Table 3** Measures collected during the experiment

Task	Measure	Description
Reading Comprehension	Reading time	Average duration covering the time from when the text passage was displayed until the “Done” button was pressed
	Response time	Average time from when the answer screen was displayed until the “Submit” button was pressed
	Scrolls	The average number of times a scroll arrow was pressed, either using an on-screen arrow or a physical arrow button on the device
	Score	The number of correct answers selected (out of 10)
	TLX	The overall TLX subjective workload score (between 0 and 100)
Word Search	Time	The average time from when the passage was displayed until a line on the screen was selected
	Score	The number of correct selections containing the target word (out of 10)
	TLX	The overall TLX subjective workload score (between 0 and 100)
Both tasks	Acceleration	Acceleration in the <i>X</i> , <i>Y</i> , and <i>Z</i> planes as well as the net acceleration, sampled at 7 Hz
Both tasks, walking condition only	Baseline walking speed, number of complete and partial laps during each scenario, number of times participant stepped on the tape during each scenario, shoe size of participant	

was the same as the first. After both tasks had been completed, participants filled out a post-task questionnaire that asked them to indicate the degree to which the various factors in the study contributed to the difficulty of the tasks.

## 2.5 Experimental measures

Table 3 describes the measures that were recorded during the experiment. Note that for the Reading Comprehension task, the task time was divided into two separate measures. This was done in order to separate the time required to understand and encode the information in the passage (reading time) from the time required to query and recall the information that had been processed during reading (response time). Among them, the acceleration data was analyzed separately and published elsewhere [37]. The data from measures only available in walking condition were not included because they cannot be used in comparing sitting and walking conditions.

## 2.6 Hypotheses

Since very little previous empirical work has been done investigating the specific effects of task type, motion, and lighting, hypotheses were generated with a broad stroke, presuming that the contextual factors would affect both tasks and all experimental measures similarly. Casual observation dictated that the effects of

motion would be greater than changes in lighting, in general. Therefore, the hypotheses for this study were as follows:

*Hypothesis 1* For the *Reading Comprehension* task, the effect of *motion* will yield *strongly significant* differences for all experimental measures.

*Hypothesis 2* For the *Reading Comprehension* task, the effect of changes in *lighting* will yield *significant* differences for all experimental measures.

*Hypothesis 3* For the *Word Search* task, the effect of *motion* will yield *strongly significant* differences for all experimental measures.

*Hypothesis 4* For the *Word Search* task, the effect of changes in *lighting* will yield *significant* differences for all experimental measures.

## 3 Results

The two tasks used in this study were not designed to be compared to each other in terms of difficulty or duration, as they represented fundamentally different types of PDA tasks. Thus, differences between them in terms of these measures were not of interest. Because of this, the results for each task will be presented and discussed separately.



In order to facilitate a more in-depth discussion, differences between conditions will be divided into three categories:

Category	$p$ value	Description
Not significant	$p > 0.05$	The conditions are considered to be equivalent in their effect on the dependent variable
Significant	$0.01 < p \leq 0.05$	Differences between conditions are very likely
Strongly significant	$p \leq 0.01$	Differences between conditions are almost certain

These categories will be designated by \* (significant) or \*\* (strongly significant) in Tables 4, 5, 7, 8, and 9.

### 3.1 Task 1: Reading Comprehension

After the data had been collected, correlation tests were run in order to determine if any of the participant demographic factors had possibly influenced the observed values of the experimental measures. Four out of the five experimental measures in the Reading Comprehension task exhibited a high degree of correlation between several demographic characteristics and the observed results. The following factors had a noticeably high degree of correlation with one or more of the experimental measures: *age*, *native language*, *dominant hand*, *PDA familiarity score*, and *response to “do you regularly read books or other printed text while walking?”* As a result of the relatively large number of demographic factors that likely had influence on the recorded measures, a repeated-measures analysis of covariance (ANCOVA) statistical analysis was used to investigate differences between the independent variables. ANCOVA has the advantage that it is able to disentangle the effects of the independent variables (in this case lighting and motion) from the effects of covariates by including them in the regression model. Each response is therefore decomposed into three parts: that which can be explained by the independent variables alone, that

which can be explained by any covariates, and that which cannot be explained by either of the above two (the error). This results in the comparison of responses that have been adjusted in magnitude to account for the effects of the covariates. ANCOVA is also robust to uncertainty about the presence of significant correlation between potential covariates and response variables, so long as there is no dependence between the treatment conditions and the covariates (which holds in this case because participants were randomly assigned to groups). For additional explanation of the appropriateness of ANCOVA for this purpose, refer to Neter et al. [38].

The benefit of using ANCOVA for this analysis is two-fold:

1. The discrepancy between the two experimental groups in terms of number of native English speakers can be accounted for by treating native language as a covariate, whereby its effect is essentially neutralized for each participant when the experimental conditions are being compared.
2. Known relevant covariates are accounted for in the regression equation, creating a less “noisy” model, thus increasing the power of the model over a standard ANOVA analysis.

Participant responses to the NASA-TLX workload assessment were not shown to be significantly correlated with the recorded demographics and were therefore analyzed using a standard ANOVA technique, after verifying that the data met the required assumptions for the test.

#### 3.1.1 Motion

The mean adjusted values for the two between-subjects (motion) conditions as well as the results of statistical comparisons are listed in Table 4. TLX scores are unadjusted because they were not shown to be correlated with participant demographics. All times are in milliseconds.

**Table 4** Adjusted values for experimental measures between motion conditions

Measure	Condition	Mean	Standard error	$F$	$p$
Reading time (ms)	Sitting	26,489	1,150	4.747	0.033*
	Walking	30,204	1,212		
Response time (ms)	Sitting	17,602	590	0.388	0.536
	Walking	18,147	621		
Score	Sitting	8.26	0.17	8.607	0.004**
	Walking	7.51	0.18		
Scrolls	Sitting	14.16	0.49	0.305	0.582
	Walking	13.67	0.52		
TLX <sup>a</sup>	Sitting	56.46	1.95	12.099	0.001**
	Walking	66.28	2.03		

An asterisk (“\*”) and double asterisks (“\*\*”) denote significant and strongly significant, respectively

<sup>a</sup>These values are the raw means and have not been adjusted for the presence of covariates

**Table 5** Adjusted values for experimental measures between lighting conditions

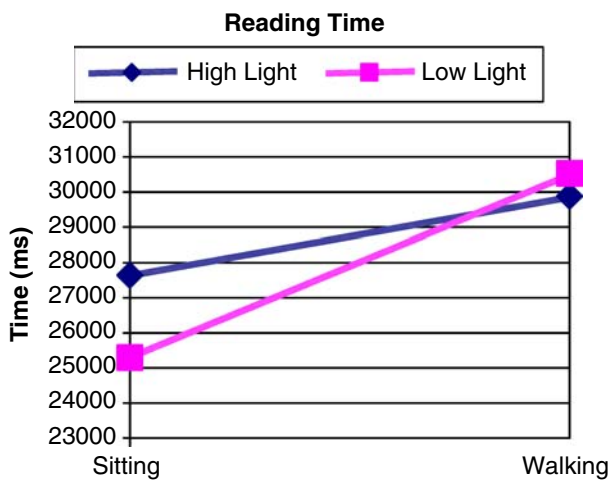
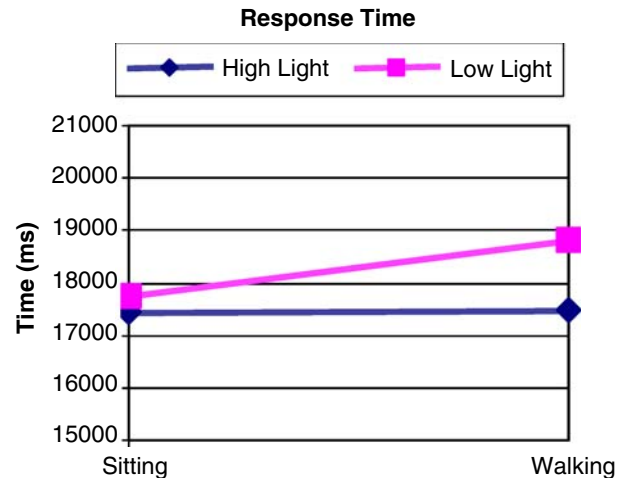
Measure	Condition	Mean	Standard error	<i>F</i>	<i>p</i>
Reading time (ms)	High-Light	28,761	898	1.053	0.308
	Low-Light	27,933	926		
Response time (ms)	High-Light	17,475	485	3.983	0.050*
	Low-Light	18,275	444		
Score	High-Light	7.93	0.17	0.127	0.723
	Low-Light	7.85	0.15		
Scrolls	High-Light	13.31	0.45	4.244	0.043*
	Low-Light	14.61	0.49		
TLX <sup>a</sup>	High-Light	60.17	1.61	4.363	0.040*
	Low-Light	62.58	1.43		

An asterisk (“\*”) denotes significant

<sup>a</sup>These values are the raw means and have not been adjusted for the presence of covariates

**Table 6** *F* and *p* values for the motion × lighting interactions

Measure	Motion condition	Lighting condition	Mean	Standard Error	<i>F</i>	<i>p</i>
Reading time (ms)	Sitting	High-Light	27,662	1,263	3.244	0.076
		Low-Light	25,318	1,302		
	Walking	High-Light	29,861	1,330		
		Low-Light	30,548	1,372		
Response time (ms)	Sitting	High-Light	17,464	682	1.575	0.213
		Low-Light	17,740	624		
	Walking	High-Light	17,485	719		
		Low-Light	18,809	657		
Score	Sitting	High-Light	8.23	0.25	0.436	0.511
		Low-Light	8.30	0.22		
	Walking	High-Light	7.64	0.26		
		Low-Light	7.39	0.23		
Scrolls	Sitting	High-Light	13.56	0.63	0.021	0.886
		Low-Light	14.77	0.69		
	Walking	High-Light	13.06	0.67		
		Low-Light	14.46	0.73		
TLX	Sitting	High-Light	56.34	2.24	3.572	0.063
		Low-Light	56.57	1.98		
	Walking	High-Light	63.99	2.32		
		Low-Light	68.58	2.06		

**Fig. 7** Reading time motion × lighting interaction**Fig. 8** Response time motion × lighting interaction

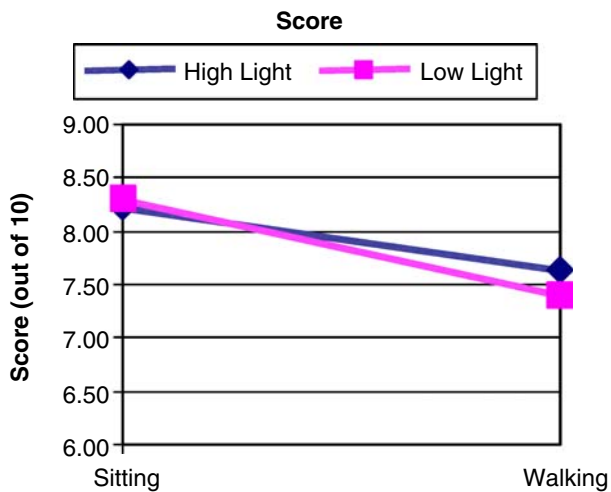


Fig. 9 Score motion × lighting interaction

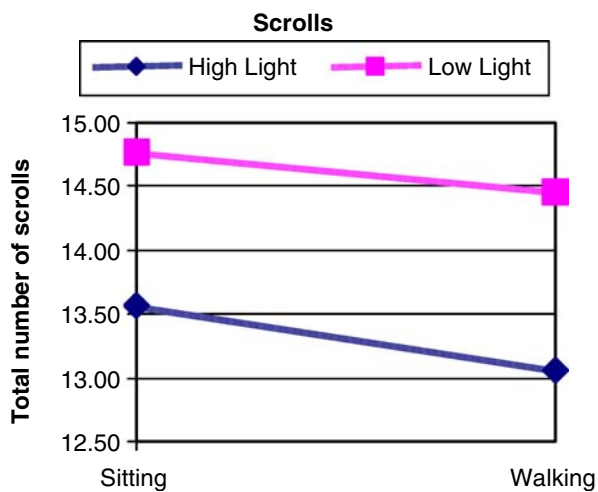


Fig. 10 Scrolls motion × lighting interaction

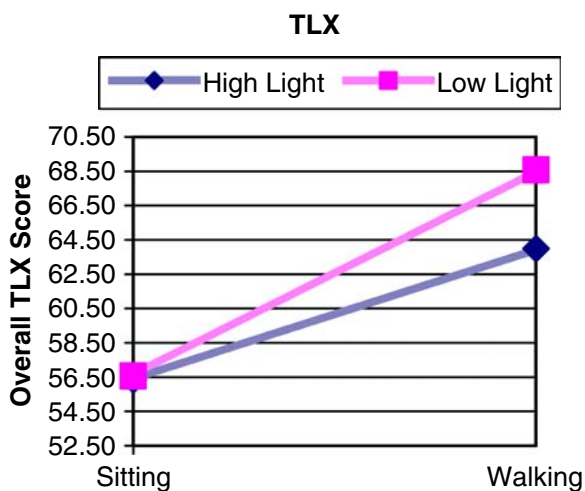


Fig. 11 TLX motion × lighting interaction

### 3.1.2 Lighting

The same data were analyzed to look at the differences in performance (in both conditions) between the High-Light and Low-Light scenarios. The mean adjusted values for the within-subjects (lighting) conditions are listed in Table 5, along with the results of the statistical analyses.

### 3.1.3 Motion × lighting interactions

In order for the statistical results in Tables 4 and 5 to be directly interpretable, there must be no indication of a significant interaction between the two independent variables. The results of the ANCOVA motion × lighting interaction are summarized in Table 6.

In order to develop a more complete picture of the effects of the two conditions on performance, the graphs in Figs. 7, 8, 9, 10, and 11 illustrate the change in performance as the independent variables were varied at each level. The four data points in each graph represent the adjusted values of the response variables for each of the four scenarios in the Reading Comprehension task.

## 3.2 Task 2: Word Search

Once again, correlations were run between the response variables and participant demographics. For the results of the Word Search task, one response variable, time, was shown to be significantly correlated with native language. Neither the score nor TLX scores were shown to be correlated with any of the demographics collected. Therefore, an ANCOVA was performed on time, while ANOVA was used for score and TLX scores.

### 3.2.1 Motion

The values of the response variables that were compared are presented in Table 7.

### 3.2.2 Lighting

The same data were analyzed to look at the differences in performance (in both conditions) between the High-Light and Low-Light scenarios. The mean adjusted values for the within-subjects (lighting) conditions are listed in Table 8, along with the results of the statistical analyses.

### 3.2.3 Motion × lighting interactions

Interestingly, two of the three 2-way interactions appeared significant, indicating that performance changed significantly with the combined effect of lighting and motion, but not due to either effect in isolation. Table 9 provides the statistical results that reveal this phenomenon.

**Table 7** Adjusted values for experimental measures between motion conditions

Measure	Condition	Mean	Standard error	<i>F</i>	<i>p</i>
Time (ms)	Sitting	3,930	151	4.945	0.029*
	Walking	4,426	159		
Score <sup>a</sup>	Sitting	9.80	0.06	4.501	0.037*
	Walking	9.60	0.06		
TLX <sup>a</sup>	Sitting	39.23	2.58	4.184	0.044*
	Walking	46.79	2.64		

An asterisk (“\*\*”) denotes significant

<sup>a</sup>These values are the raw means and have not been adjusted for the presence of covariates

**Table 8** Adjusted values for experimental measures between lighting conditions

Measure	Condition	Mean	Standard error	<i>F</i>	<i>p</i>
Time (ms)	High-Light	4,003	130	6.728	0.011*
	Low-Light	4,352	124		
Score <sup>a</sup>	High-Light	9.74	0.06	0.821	0.368
	Low-Light	9.66	0.06		
TLX <sup>a</sup>	High-Light	41.68	2.05	3.050	0.085
	Low-Light	44.33	1.94		

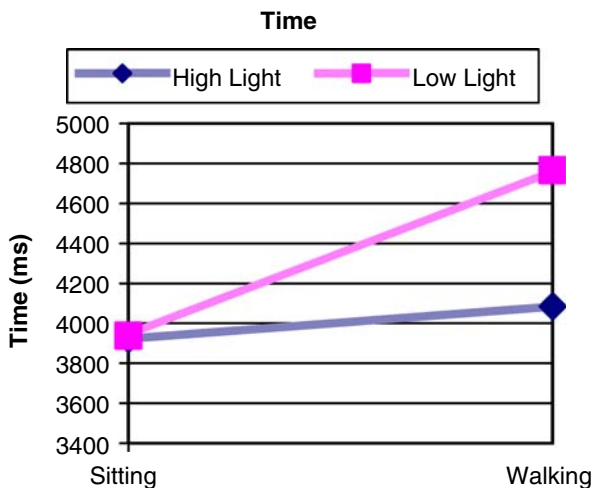
An asterisk (“\*\*”) denotes significant

<sup>a</sup>These values are the raw means and have not been adjusted for the presence of covariates

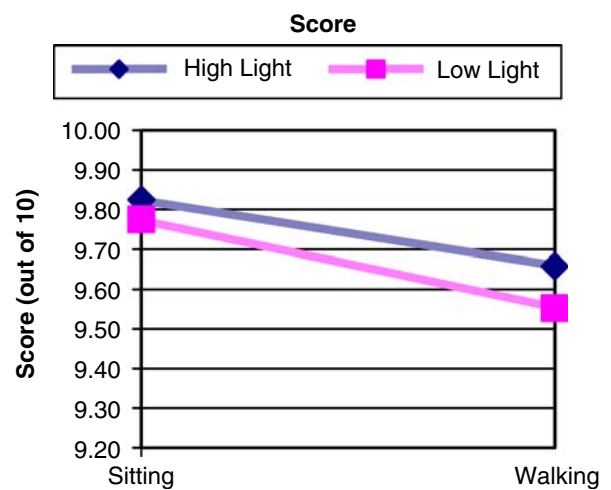
**Table 9** *F* and *p* values for the motion × lighting interactions

Measure	Motion condition	Lighting condition	Mean	Standard error	<i>F</i>	<i>p</i>
Time (ms)	Sitting	High-Light	3,921	182	5.733	0.019*
		Low-Light	3,939	173		
	Walking	High-Light	4,085	192		
		Low-Light	4,766	182		
Score	Sitting	High-Light	9.82	0.08	0.104	0.748
		Low-Light	9.77	0.09		
	Walking	High-Light	9.65	0.08		
		Low-Light	9.55	0.09		
TLX	Sitting	High-Light	39.88	2.86	6.831	0.011*
		Low-Light	38.57	2.71		
	Walking	High-Light	43.48	2.94		
		Low-Light	50.10	2.78		

An asterisk (“\*\*”) denotes significant



**Fig. 12** Time motion × lighting interaction



**Fig. 13** Score motion × lighting interaction



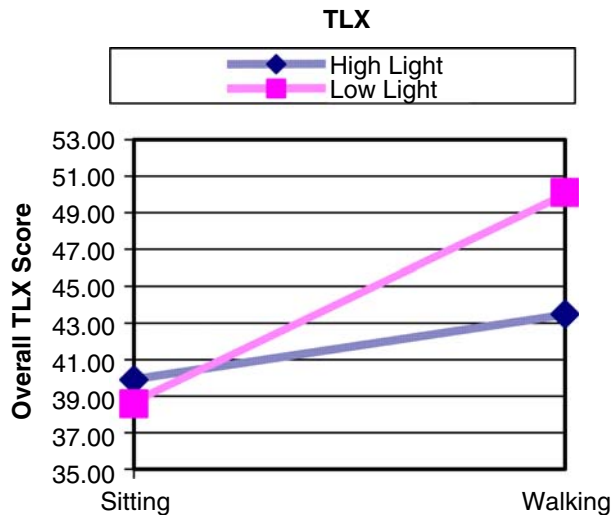


Fig. 14 TLX motion × lighting interaction

The graphs of the interactions, shown in Figs. 12, 13, and 14, clarify this result.

For the time and TLX measures, nonparallel lines can be seen, indicating the statistical significance of the interaction, where the effect of one factor is not consistent across the other. In the cases of time and TLX, changes in lighting had little effect in the sitting condition, but had a noticeable effect in the walking condition. Similarly, changing from sitting to walking had only a small effect in the High-Light condition, but a large effect in the Low-Light condition. Implications of these results will be discussed in the proceeding section.

#### 4 Discussion

First of all, it is important to remember that this study only looked at two contextual factors at two levels for two tasks and the results are not intended to be representative of all tasks under all conditions. However, the results that were generated are rich with valuable information and have some degree of generalizability because the factors were carefully chosen to be representative of aspects of context encountered on a daily basis. Because context is so complex, it is important to have some degree of control when investigating its effects, therefore incremental advancements in the understanding of context may yield more benefit in the long term than attempts to quantify all effects at once.

*Hypothesis 1:* For the *Reading Comprehension* task, the effect of *motion* will yield *strongly significant* differences for all experimental measures.

This hypothesis was partially confirmed. Score and TLX were shown to be strongly significant ( $p=0.004$  and  $0.001$ , respectively) between motion conditions, while reading time was significant ( $p=0.033$ ). Interest-

ingly, response time and scrolls were clearly nonsignificant ( $p=0.536$  and  $0.582$ , respectively). The lack of differences in response time could reveal that users were unable to process the text passages deeply, even though they took more time on them. This could indicate more difficulty in information encoding than information retrieving. Presumably, participants would have spent more time answering the questions in the walking condition than in the sitting condition if they felt that the necessary information was available to them, but was more difficult to access. It could be the case that, even though participants spent more time reading the passages and trying to encode them in the walking condition, the process of encoding was hindered by their motion. Thus, differences in the score measure between motion conditions were caused by inefficient processing of information during the reading phase. The lack of differences in the scrolls measure could indicate that participants were not less efficient in their reading strategy during the walking condition, because they did not use the scroll buttons more frequently.

*Hypothesis 2:* For the *Reading Comprehension* task, the effect of changes in *lighting* will yield *significant* differences for all experimental measures.

This hypothesis was also partially confirmed. Similar to the results for the effect of motion, lighting yielded significant differences for some measures (response time,  $p=0.050$ ; scrolls,  $p=0.043$ ; TLX,  $p=0.040$ ) and nonsignificant differences for others (reading time,  $p=0.308$ , score,  $p=0.723$ ). Interestingly, the measures that showed up as significant were, with the exception of TLX, the nonsignificant measures in the motion results. This indicates that changes in lighting affected users in a fundamentally different manner than changes in motion. Apparently, lighting affected users in a slightly more superficial way, leading to decreased reading efficiency (noted by more instances of scrolling) and response selection speed, but not accuracy. It is important to note that all noteworthy differences were significant and not highly significant, indicating that the changes in motion are able to influence user behavior in a more dramatic way than changes in lighting.

*Hypothesis 3:* For the *Word Search* task, the effect of *motion* will yield *strongly significant* differences for all experimental measures.

This hypothesis was confirmed, but not as strongly as anticipated. This was the only instance where the effects of a contextual factor were fairly consistent across all measures (time,  $p=0.029$ ; score,  $p=0.037$ ; TLX,  $p=0.044$ ). However, since there were fewer measures recorded for this task, some instances of similar behavior between the two conditions could have been overlooked. None of the measures were shown to be strongly significant between the two motion conditions, however. This is likely due to the fact that less deep processing was required and because the task was slightly less difficult than the *Reading Comprehension* task. This could

indicate that relatively shallow processing like word recognition is marginally affected by differences in motion; however, a more difficult task might have shown more strongly significant differences, thus disproving this conjecture.

*Hypothesis 4:* For the *Word Search* task, the effect of changes in *lighting* will yield *significant* differences for all experimental measures.

This hypothesis was partially confirmed, as one of the three measures (time;  $p = 0.011$ ) was significant. As in the Reading Comprehension task, score was unaffected by changing light conditions ( $p = 0.368$ ). TLX was also not significant ( $p = 0.085$ ), but close enough to be noted because of the subjectivity of the measure, as it is noisier in general than the other measures. This is of interest because it could provide insight into participants' prioritization during the task. It is possible that they put accurate performance as their primary objective, and compromised their time and effort in order to perform accurately. This indicates the flexibility of users to adapt their behavior and make choices when faced with limiting environmental conditions.

Overall, the overarching hypothesis that the independent variables would influence all experimental measures similarly was generally not supported, as some measures, particularly for the Reading Comprehension task, were very significant, while others were quite nonsignificant. This result is compelling, as it indicates that the way in which users' behavior is affected by changes in context is not uniform. The other overarching hypothesis that users would be impacted similarly in both tasks was partially supported, as measures of time, score, and workload were significantly different between motion conditions. However, there were some discrepancies in the results for the lighting effects, as well as the degree of significance of the interactions. This is interesting because it indicates some effect of task type on user behavior even when the device and scenarios are the same.

Additionally, the interactions between the contextual factors, which were not addressed in the initial hypotheses, presented some of the most interesting results. Time and TLX showed significant differences ( $p = 0.019$  and  $0.011$ , respectively) for the motion  $\times$  lighting interactions for the Word Search task, while the time and workload interactions were nearing significance ( $p = 0.076$  and  $0.063$ , respectively) for the Reading Comprehension task. When one considers the relatively controlled nature of this study, where most other contextual factors were held constant, the implications of this result for true real world conditions are enlightening. It should be a fairly safe assumption to conclude that the interaction effects would be even more dramatic when the number of variable contextual factors is increased, as in a typical real world mobile interaction scenario. This strongly indicates that mobile device evaluation in a static, seated, environment is likely to elicit far different behavior from users than they would exhibit in a real world situation.

While this study only examined a small number of contextual factors, the results clearly indicate that common contextual variation (switching from sitting to walking, high light to low light, one task to another) has a clear, but diverse effect on the way in which users interact with a mobile device. Similar research looking at other contextual factors, or other levels of similar factors, would yield much needed empirically based insight into human behavior with mobile devices and allow context to be modeled and designed for much more appropriately.

---

## 5 Conclusions

Mobile computing devices can be used in myriad environments where a desktop computer would never be found; in the hands of a standing passenger on a bus, with a doctor who is diagnosing a patient while they are being transported, as a tour guide in a museum, for example. The situational context surrounding the interaction with a mobile device by a human user can greatly influence the user's behavior and their abilities relating to their mobile device. Even if a mobile device user is not moving, the fact that a mobile device can be taken practically anywhere dictates that it is subject to a wide variety of constraints relevant to the user. In order to overcome the challenge of designing for mobility and context-rich environments, new paradigms, such as the concept of SIID which draws analogies between the impairments generated by contextual factors and those occurring as a result of disabilities, should be investigated as well. For example, existing knowledge about effective strategies for designing for persons with disabilities can likely be leveraged to assist with designing for mobile device users.

In order for mobile computing to meet its grand expectations, it is imperative that contexts of use, particularly as they apply to the user, be considered and that mobile devices be designed with these contexts in mind. Additionally, this study has indicated that creating contextually rich scenarios for mobile device evaluation is not excessively complicated, primarily because context is all around us. And, with the ever-increasing availability of inexpensive, integrated, or compact sensors for mobile devices, measuring and modeling the effects of context is becoming less and less intimidating.

It is important that more research be conducted to investigate the ways in which changes in context impact user behavior because, as this study has shown, context does not affect people in a uniform way. Beyond that, people have a choice as to how they react in context-rich environments, and their behavior is often dictated by their own priorities as well as their abilities. The way in which users allocate available cognitive and physical resources when using mobile devices is very important. Users may be able to maintain adequate performance levels on a mobile device, but the expense may be too costly in certain situations. When available attentional

resources are scarce, such as when driving or performing other safety-critical, focused tasks, or in time-sensitive situations, consideration of context and the way people manage multiple task demands is especially important. Context cannot be defined by independently considering specific contextual components and adding them together. Context is, by nature, a multifaceted construct, and the ways in which contextual factors interact, combine, and consolidate need to be studied further by both researchers and practitioners.

This project was limited to the domain of pen-based mobile computing in low-stress environments, yet the results indicate that context is a rich, nuanced, and variable condition. The authors believe that context is relevant and applicable in almost every situation and that investigations of context will be fruitful in any domain where a user has a specific goal that they are working toward, yet has multiple variables vying for their attention. These investigations should be catered to the domain being studied and should be designed to mirror realistic situations as closely as possible, while still retaining experimental control.

Some specific areas that deserve additional attention include nomadic interactions, such as those described in the current article, as well as the potential influence of variable lighting. Variable levels of noise may also prove worthy of investigation, especially in the context of speech-based interactions. Extensive research has focused on improving speech recognition algorithms, making them more robust in the context of noise, and on developing hardware-based solutions such as noise canceling microphones. However, little has been reported with regard to user interactions with speech-based solutions under realistic conditions that include variable levels of noise. Ultimately, effective solutions to the challenges introduced by varying context will require a combination of hardware (e.g., noise canceling microphones), software (e.g., algorithms to stabilize stylus inputs when users are on the move), and careful design.

**Acknowledgements** The authors are greatly indebted to Paula Edwards, V. Kathlene Leonard, Thitima Kongnakorn, and Kevin Moloney for their crucial assistance and support. This material is based upon work supported by the National Science Foundation (NSF) under Grant No. IIS-0121570. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

## References

- Brewster S, Leplâtre G, Crease MG (1998) Using non-speech sounds in mobile computing devices. In: First workshop on human computer interaction with mobile devices, Glasgow, Scotland
- Johnson P (1998) Usability and mobility: interactions on the move. In: First workshop on human computer interaction with mobile devices, Glasgow, Scotland
- Kristoffersen S, Ljungberg F (1999) Designing interaction styles for a mobile use context. In: HUC'99. Springer, Berlin Heidelberg New York
- Oulasvirta A (2004) Finding meaningful uses for context-aware technologies: the humanistic research strategy. *CHI Lett* 6:247–254
- Beale R, Lonsdale P (2004) Mobile context aware systems: the intelligence to support tasks and effectively utilise resources. In: Sixth international mobile HCI conference, Glasgow, Scotland. Springer, Berlin Heidelberg New York
- Dourish P (2004) What we talk about when we talk about context. *PUC* 8:19–30
- Marmasse N, Schmandt C (2000) Location-aware information delivery with ComMotion. In: Handheld and ubiquitous computing, Bristol, UK. Springer, Berlin Heidelberg New York
- Dey AK, Abowd GD, Salber D (1999) A context-based infrastructure for smart environments. In: International workshop on managing interactions in smart environments (MANSE '99)
- Jang S-I, Kim J-H, Ramakrishna RS (2001) Framework for building mobile context-aware applications. In: *Human.Society@Internet*. Springer, Berlin Heidelberg New York
- Hess CK, Campbell RH (2003) An application of a context-aware file system. *PUC* 7:339–352
- Van der Meer S, Arbanowski S, Popescu-Zeletin R (1999) A platform for environment-aware applications. In: International symposium on handheld and ubiquitous computing
- Sears A et al (2003) When computers fade... pervasive computing and situationally-induced impairments and disabilities. In: Stephanidis C (ed) *Human-computer interaction: theory and practice (Part II)*. Lawrence Erlbaum Associates, Hillsdale, pp 1298–1302
- Kjeldskov J, Graham CA (2003) Review of mobile HCI research methods. In: Fifth international symposium on human computer interaction with mobile devices and services, Udine, Finland
- Brewster S (2002) Overcoming the lack of screen space on mobile computers. *PUC* 6:188–205
- Hinckley K et al (2000) Sensing techniques for mobile interaction. In: Proceedings of the 13th annual ACM symposium on user interface software and technology, San Diego, California
- Schmidt A, Beigl M, Gellersen HW (1999) There is more to context than location. *Comput Graph* 23:893–901
- Gellersen HW, Schmidt A, Beigl M (2000) Adding some smartness to devices and everyday things. In: Workshop on mobile computing systems and applications, Monterey, California. IEEE Computer Society Press, Los Alamitos
- Barkhuus L (2003) Context information vs. sensor information: a model for categorizing context in context-aware mobile computing. In: Collaborative technologies and systems conference, San Diego, CA
- Randell C, Muller H (2000) Context awareness by analysing accelerometer data. In: Fourth international symposium on wearable computers, Atlanta, GA. IEEE Computer Society, Los Alamitos
- Rodden T, Chervest K, Davies N (1998) Exploiting context in HCI design for mobile systems. In: First workshop on human computer interaction with mobile devices. University of Glasgow, United Kingdom
- Lieberman H, Selker T (2000) Out of context: computer systems that adapt to, and learn from, context. *IBM Syst J* 39:617–632
- Korpipää P et al (2003) Bayesian approach to sensor-based context awareness. *PUC* 7:113–124
- Intille S et al (2004) Acquiring in situ training data for context-aware ubiquitous computing applications. *CHI Lett* 6:1–8
- Bellotti V, Edwards K (2001) Intelligibility and accountability: human considerations in context aware systems. *HCI* 16:193–212
- Kristoffersen S, Ljungberg F (1999) Mobile use of IT. In: Twenty-second information systems research seminar in Scandinavia
- Bristow HW et al (2004) Defining and evaluating context for wearable computing. *Int J Hum-Comput Stud* 60:798–819
- Dunlop M, Brewster S (2002) The challenge of mobile devices for human computer interaction. *PUC* 6:235–236

28. Kjeldskov J et al (2004) Is it worth the hassle? Exploring the added value of evaluating the usability of context-aware mobile systems in the field. In: Sixth international mobile HCI 2004 conference, Glasgow, Scotland. Lecture notes in computer science. Springer, Berlin Heidelberg New York
29. Kjeldskov J, Stage J (2004) New techniques for usability evaluation of mobile systems. *Int J Hum-Comput Stud* 60:599–620
30. Mustonen T, Olkkonen M, Hakkinen J (2004) Examining mobile phone text legibility while walking. In: Proceedings of CHI 2004. ACM Press, Vienna, pp 1243–1246
31. Pascoe J, Ryan N, Morse D (2000) Using while moving: HCI issues in fieldwork environments. *TOCHI* 7:417–437
32. Barnard L et al (2005) An empirical comparison of use-in-motion evaluation scenarios for mobile computing devices. *Int J Hum-Comput Stud* 62:487–520
33. Open eBook Forum (2003), Q3 2003 eBook Statistics
34. Learning express (1999) 501 Reading comprehension questions. LearningExpress, New York, NY
35. Vernier Software & Technology, Accelerometers
36. Hart SG, Staveland LE (1988) Development of NASA-TLX (task load index): results of empirical and theoretical research. In: Meshkati N (ed) Human mental workload. Elsevier, Amsterdam, pp 139–183
37. Yi JS et al (2005) Context awareness via a single device-attached accelerometer during mobile computing. In: 6th international mobile HCI 2005 conference, Salzburg, Austria, 19–22 September 2005, pp 303–306
38. Neter J et al (1996) Applied linear statistical models. WCB McGraw-Hill, Boston