Interaction in 4-Second Bursts: The Fragmented Nature of Attentional Resources in Mobile HCI

Antti Oulasvirta [1]**, Sakari Tamminen** [2]**, Virpi Roto** [3]**, and Jaana Kuorelahti** [2]
[1] Helsinki Institute for Information Technology
[2] Helsinki University of Technology
[3] Nokia Research Center
Helsinki, Finland
antti.oulasvirta@hiit.fi, {sakari.tamminen, jaana.kuorelahti}@hut.fi, virpi.roto@nokia.com

ABSTRACT

When on the move, cognitive resources are reserved partly for passively monitoring and reacting to contexts and events, and partly for actively constructing them. The Resource Competition Framework (RCF), building on the Multiple Resources Theory, explains how psychosocial tasks typical of mobile situations compete for cognitive resources and then suggests that this leads to the depletion of resources for task interaction and eventually results in the breakdown of fluent interaction. RCF predictions were tested in a semi-naturalistic field study measuring attention during the performance of assigned Web search tasks on mobile phone while moving through nine varied but typical urban situations. Notably, we discovered up to eight-fold differentials between micro-level measurements of attentional resource fragmentation, for example from spans of over 16 seconds in a laboratory condition dropping to bursts of just a few seconds in difficult mobile situations. By calibrating perceptual sampling, reducing resource usage for tasks of secondary importance, and resisting the impulse to switch tasks before finalization, participants compensated for the resource depletion. The findings are compared to previous studies in office contexts. The work is valuable in many areas of HCI dealing with mobility.

ACM Classification Keywords: H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

Keywords: Attention; cognition; mobile HCI; semi-naturalistic field study; context; multitasking; mobile browsers; multi-modal interfaces; interruptions

INTRODUCTION

People seem to have tremendous capabilities for utilizing mobile devices in various innovative and fulfilling ways while on the move. However, there are times when the fluency of interaction breaks down dramatically. We all have experiences where we have to slow down, to postpone, or to stop interaction with a device entirely because of a cognitively taxing situation. And we sometimes have to invent novel ways of interaction or workarounds on the spot. From the perspective of interaction, then, being mobile is cognitively costly.

In this paper, we investigate how mobile contexts cause such shortages of cognitive resources. Our starting point is that there are typical tasks related to mobile contexts (e.g., a crowded bus context and the task of managing personal space) and these tasks differ in how they tax attention. Building on the Multiple Resources Theory of attention [38], we propose the Resource Competition Framework that relates mobile task demands to user's cognitive resources. It predicts differences in the availability of attention for interaction among intuitively equal contexts. The predictions are tested in a semi-naturalistic field study where minicams are utilized to record attention on the move. Finally, we link the results to several topical research questions in mobile HCI.

Previous Research

In a pioneering study, Kristoffersen and Ljungberg [16] explained that mobile devices reserve our physical and attentional capabilities (e.g., one hand for holding the phone) from other tasks required for mobility. They observed that mobile workers need to "make a place" for the device in a taxing situation (e.g., drivers may need to stop the car to release enough resources to operate a phone). Similar observations were made by Pascoe *et al.* [25], and by Lumsden and Brewster [18] who related them to a wider perspective by noting that there is often a conflict between mobility and mobile HCI: Interacting with a mobile device, as a task, competes for the same limited resources that we need to safely navigate through the environment [32].

The relevance of social contexts to the use of attentional resources was identified when researchers began to examine mobility from a broader perspective. Perry *et al.* [27] stressed that mobility involves uncertainty and unpredictability, which demands "planful opportunism". Similarly, our ethnomethodologically oriented mobility study revealed that the sociality of mobility demands not only attention but also active participation [34]. Five observations were made:

CHI 2005, April 2–7, 2005, Portland, Oregon, USA.

(1) Mobility is often bound to social goals related to collaboratively produced organizations of time-space of near future (meetings, deadlines etc.); (2) Situated acts are embedded within planned ones—dropping by, ad hoc meetings, and other forms of sidestepping require planful opportunism and flexibility from other mobility plans—particularly navigation [27,33]; (3) Personal spaces must be actively claimed and held by socially recognizable acts, because mobile contexts are typically public and dynamic rather than private and stable; (4) Aspects of mobility impose various social and cognitive restrictions on multitasking; (5) Finally, and primary significance to the present paper, different temporal tensions ("fluctuations of importance of time in relation to space"; e.g., hurrying, decelerating, waiting) are related to different cognitive and social demands for action. Subsequently, we made some initial observations on how task demands posed by mobile contexts "create" these tensions [21].

These reflections provided the impetus for the current work.

Mobile Contexts as Interactional Achievements

Our starting point is that human action has its basic contextual parameters [12]: (1) the setting (social and spatial framework), (2) behavioral environment, (3) linguistic, and (4) extrasituational context. The actor must orient to these contextual parameters, and this orientation in turn has the potential of becoming context-creating and -renewing for the self and for others. For example, visiting a church requires the visitor to orient to the social (e.g., a wedding), spatial (chairs, walls), behavioral (gestures of people near by), language (talk), and extrasituational (what is known in advance about the particular wedding) contexts. By taking action, or withholding it, the visitor can either renew or sustain the present context. The fundamental idea, then, is that people, through a set of various purpose-oriented and streamlined "ethnomethods", construct mobile contexts. This view comes close to the interactional view of context [9] inspired by ethnomethodology. An opposing view is the representational view where contexts are regarded as observer-independent entities ready to be sensed, represented, and acted upon by an agent.

We believe that adopting the interactional view is essential for understanding why mobile situations yield cognitive costs. For example, waiting for a metro to arrive is not simply about sitting idly with all cognitive resources free for time killing activities, but calls for action: estimating when the metro arrives, moving to a position where it can be perceived, continuously interpreting auditory sense data, monitoring how personal space is perhaps intruded by by passers, occasionally glancing the environment to see if the metro is coming etc. [34] These *mobility tasks* compete for cognitive resources with other tasks, including *mobile HCI tasks*, the most important tasks given higher priority and "left-over" resources being re-distributed.

This paper proposes an elaborated view on how this occurs.

THE RESOURCE COMPETITION FRAMEWORK (RCF)

Successful context construction in mobility is a complex multitasking achievement, requiring careful orchestration of the planning, timing, execution, monitoring, and control of tasks. One must switch back and forth between tasks and external sources, temporarily leaving the switched-from tasks on hold or slowing them down. The key to this achievement is controlled selection and division of attention. RCF attempts to describe the necessary cognitive faculties underlying this achievement in mobile HCI.

As an analytical tool, RCF builds on three cognitively oriented traditions. First, RCF borrows from the tradition of task analysis. Both analyze the tasks "required" of a person in a situation, decompose them to their components, and identify related mental requirements. Task analysis has been used extensively in studies of pilots, drivers, and control room operators, but not yet of mobility. Second, RCF also comes close to the thinking in the cognitive modeling tradition in HCI (e.g., GOMS, EPIC, ACT-R; [4]). However, it does not aim to model the interplay of cognitive faculties, but to identify and explicate the relevant cognitive resources and their properties. Third, regarding cognitive psychology, RCF borrows from the work of Navon [19] on the resource approach to attention and particularly from the work of Wickens on Multiple Resource Theory [38]. RCF is custom-tailored for mobility and is thus less fine-grained and lacks the notion of representation codes. Recent findings have also challenged some of the assumptions of Wickens, in particular the fixed capacity assumption [39]. Nevertheless, keeping in mind the practical orientation of this paper, the idealization can be kept.

Assumptions

In what follows, we elaborate the notion of resource competition. Four assumptions about resources are made:

- *Functional modularity*. The cognitive system is divided into functionally separate subsystems operating on different representations at different levels of processing.
- *Parallel module operation*. Cognitive modules can operate to a large extent in parallel, independent of each other, although structural interference can occur.
- *Limited capacity*. Cognitive modules are limited in capacity, with respect to time or content (e.g., perception). Capacities can serve various tasks concurrently, resulting in the desired level of performance for each, provided that the total capacity has not been exceeded. [6].
- *Serial central operation*. Central co-ordination of modules (e.g., monitoring, manipulation, inter-module information transfer, and response selection) is serial [1,26].

Second, five assumptions concern "*competition*":

- *Multitasking*. At any time, the cognitive machinery performs multiple tasks.
- *Resource pooling*. The operated tasks may pool the limited resources [19].
- *Task differences*. Different tasks require different quantities of different cognitive resources. Automatization and

#	Cognitive faculty	Provided resource	Limitations important in mobility and interaction
1	Motor control	Sequencing, timing, control, and finalization of motor actions	Fluent execution dependent on learned procedural skills
	(hands)	Manipulation of physical objects	Reach, capacity, strength, dexterity and accuracy
	(legs)	Positioning and moving the self	Limited speed, stamina
2	Sensation	Intake of external stimuli	Acuity, accuracy
3	Perception	Organization of sense data	Uniformity of perception
4	Attention (visual) (auditory) (motor)	Search, selection, and integration in and across modality-bound representations	Limited spatial span, limited object span
5	Central executive	Meta-level control of cognitive operations -selection, inhibition, updating, shifting	Serial processing
6	Working memory (visual) (auditory) (motor)	Retention of previously attended information	Short retention span, small capacity
7	Prospective memory	Proactive control and timing of actions	Susceptibility to interference
8	Episodic memory	Mental "time travel"	Distortions, slow and effortful access
9	Semantic memory	Fact knowledge	Susceptibility to interference
10	(Conscious) Thought	Conscious manipulation of Working memory materials -abstraction, inference, reasoning, problem-solving	Uniformity of thought, reliance on heuristics

Table 1. Cognitive faculties relevant in mobile HCI, their provided resources and known limitations.

skills can modify the task's need for resources [10].

- *Preferential resource sharing*. Resources are not allocated equally but hierarchically according to intrinsic motivational structures like needs and goals [8].
- *Resource-depletion penalty*. Tasks that do not receive sufficient resources are slowed-down, postponed, put on hold, or terminated [1,10,37].

Consequently, two kinds of multi-task interference can occur, both leading to a resource-depletion penalty: (1) *structural interference* occurs when two or more tasks compete for limited resources of a peripheral system; (2) *capacity interference* occurs when the total central processing capacity has been exceeded by multiple concurrent tasks. It is also assumed that *compensatory strategies* are adopted to overcome negative consequences of depletion [10]; however, RCF cannot predict what these strategies are.

Cognitive Analysis of Mobility Tasks

In Table 1, we suggest (1) central cognitive faculties and (2) their respective resources and limitations that have practical significance for mobility and HCI tasks [23]. In Table 2, we list and analyze resources of typical *mobility tasks* vs. *mobile HCI tasks*. The analysis in Table 2 is based partially on existing literature, especially [17] and our previous analyses [21-23,34]. Although RCF is more general, we here focus particularly on the deployment of visual gaze patterns and some aspects of motor control.

Mobility tasks	Mobile HCI tasks (Web search)
A. Talking (2,3,4,5,6,7,8,9,10)	J. Typing information (1,2,3,4,5,6)
B. Walking (1,2,3,4,7)	K. Deciding on a path (4,5,6,8,9,10)
C. Waiting (2,3,4,5,7)	L. Waiting for loading (2,3,4,7)
D. Way finding (4,5,6,7,8,9,10)	M. Searching from display (2,3,4,6)
E. Sidestepping (5,6,7,9,10)	N. Exec. a navig. action (1,2,3,4,5)
F. Planning routes (5,6,7,8,9,10)	
G. Avoiding collisions (1,2,3,4)	
H. Estimating time-to-target (4,5,6,7)	
I. Controlling personal space (1,2,3,4,5)	
Mobility tasks in some mobile situations in the study	
Hurrying through a busy street (B,D,F,G,H)	
Laboratory (I, if experimenter present)	
Traveling in a bus/metro (C,D,H,I)	
Chatting and drinking coffee in a café (A,C,E,I)	
Standing in a busy spot in railway station (G,H,I)	
Waiting for a metro to come on a platform (C,H,I)	

Table 2. Examples of cognitive demands in mobility vs. HCI tasks. Numbers in parentheses refer to Table 1.

Hypotheses

In the following, we expand on and make hypotheses and expand on of some common situations included in our field study. As will be shown, common mobile situations can be highly variable in terms of attentional resource competition. (See also the bottom part of Table 2.)

Walking through a busy street to a bus stop is a taxing task requiring planning a route, managing time-to-target, and walking while at the same time taking care of safety (avoiding collisions; e.g., being hit by a car). It is important to note that perceiving the bus stop, navigating, and handling personal space all compete for attention. Perceptual sampling of the environment must take place frequently.

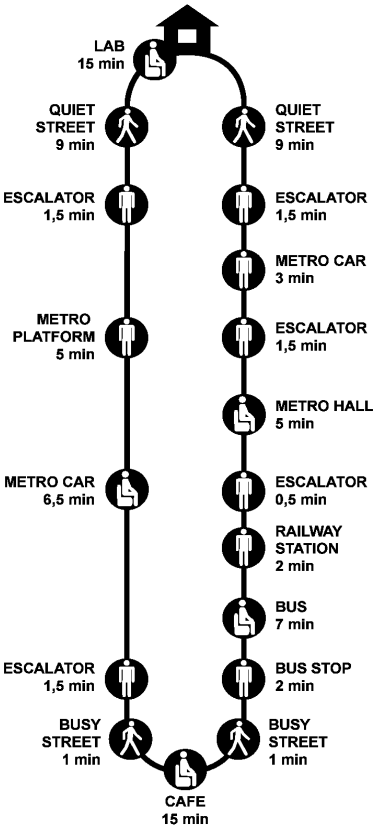
Traveling on an escalator requires less motor control, although body posture must be monitored and the right hand is usually reserved for holding on. By contrast, visual attention is more taxed: a correct standing position must be chosen, personal space must be monitored (as by passers move close), and the rapidly approaching end signal must be perceived (i.e., in order to know when to step off). Visual sampling is not as frequent as in the previous case.

Conversing in a café, on the other hand, is a task that does not require body movement except for what is needed to sip the coffee, gestures to support the talk, and monitoring social surroundings and personal space. It does require attending to the other person, inferring, making sense, and finally responding to her in a turn-taking manner [31]. It demands more of our higher-level cognitive faculties such as long-term memory and thought than visual attention.

It is interesting to compare these three to a *laboratory* situation. This context, free of disruptions and socio-cognitive demands, should involve least competition for attention.

Our exercise in the analysis of cognitive demands of Web search tasks (Table 2) implies that visual and motor resources are there the central ones. *Competition for attention* should be most pronounced in such mobile situations that also demand visual and motor capacities. Therefore:

Busy street > Escalator > Café > Laboratory.



9 Situations on the Route

Laboratory. Quiet, small laboratory room, no windows

Quiet street. Long, quiet street with clear visibility

Escalator to the metro platform/hall down/upstairs

Metro platform. Participant gets onto the next metro

Metro car. Participant's task is to sit down and get off at an instructed stop (Sörnäinen)

Busy street. A narrow, curvy street. The task is to find a way to a bus stop and stop a given bus

Caféteria. The task is to eat sandwich and drink coffee/tea. If no natural discussion emerges with other people present in the café, the experimenter initiates small talk about issues unrelated to the present study

Bus. The task is to sit in a crowded bus, stop it, and get out at an instructed stop

Railway station. The task is to stands in the middle of the hall, lots of people pass by

Figure 1. The route consisted of several places and transitory places between them. All route locations are in Helsinki.

Another hypothesis we tentatively examine here is that mobility tasks, more tightly related to social and personal well being, are at a *higher* level in the goal hierarchy than HCI tasks. For example, avoiding getting lost, slipping, or violating personal spaces should be more important than finding a Web page, even in an experiment were the participants know that their HCI performance is evaluated. To test this assumption, we include a condition in the study where participants are instructed to hurry with doing the HCI task. Because hurrying per se is not a task requiring any capacity (except keeping the instruction in mind), but only an instruction to share more resources to the HCI task, it should not lead to resources being depleted from mobility tasks, if they are more important. Hurrying is compared to waiting instruction that by contrast would be expected to release resources from the HCI task to mobility tasks by lowering its preference in the face of competition. In the baseline condition, time pressure is only implicit. Taken together, if mobility tasks are more fundamental to personal well-being than HCI tasks, we would expect hurrying in the HCI task to have *no* significant effect on the self-deployment of attention. On the other hand, we would expect waiting to release more resources for attention the environment.

We are also interested in observing how resource-depletion is compensated by adopting *attentional strategies*, for example slowing down or postponing secondary tasks.

METHOD

In the study, participants' behavior, action, and context were recorded during Web search taking place in various mobile situations. Because uncontrollable aspects of the environment and participant behavior, special arrangements were needed in terms of both study design and apparatus.

In operationalizing resource competition, our key measure was the deployment of visual gaze during Web search. We focused in particular on what happens during the subtask of page loading, instead of other more interactive tasks, for the following reasons. First, in order to perform the task quickly, the participant has to know when the page has loaded. However, as attending page loading is not as crucial to the completion of the task as for example reading the page, we expect more off-task looking and thus more sensitivity to gaze patterns. Second, as loading times vary considerably, and are thus both uncontrollable and unpredictable to the participant, he/she cannot adjust attentional strategies to estimates of loading durations, but the loading progress has to be attended to ensure rapid resumption of the task. Finally, page loading behavior is less influenced by individual interaction strategies arising from different search goals, Web page designs, and situational factors. We believe that page loadings provide a sensitive measure allowing for gathering more homogenous data.

Participants

28 subjects participated in the study. 15 of them were 20–26 and 13 of them 41–47 years old. They were experienced in using mobile phones (*M*=7.5 years, *SD*=3.5) and browsing the Web with PC (*M*=6.7 years, *SD*=2.3), and in knowing the Helsinki area (for *M*=24.1 years, *SD*=14.8) and using its public transportation system (*M*=6.2 years, *SD*=6.3). None of them had prior experience with mobile browsers.

Design

Participants in both age groups were randomly assigned to (1) a route direction (normal or reverse) and (2) task order (normal or reverse). These counterbalanced sets were created in order to tackle order effects. Instructed Time Pressure (ITP, see below) conditions were assigned to natural reference situations (some ITPs could not be assigned to certain situations, e.g. wait-ITP to walking situations). With repetitions of the situations (e.g., there were several escalator, metro, and walking situations) within a set, a different ITP was administered each time if possible. Thus, the order of ITPs was only partially counterbalanced; ITPs could not be entirely separated from the nine locations.

Materials

The Web search tasks consisted of retrieving a piece of information from a given website and reporting that to the experimenter (e.g., "Find your favorite item from today's menu at the University restaurant"). Well-known leisure time related sites were selected, most of them by or about commercial or public services in the local region. At least one interaction step had to take place to perform the task.

No text input was required. Altogether 25 tasks were created, all in Finnish.

In a study like this, the route itself is an inherent part of both stimulus materials and procedure. The route consisted of several places in the Helsinki city center. Locations, situations, transportation, and times are given in Figure 1.

Training and Procedure

Before the trials, the experimenter greeted the participant, committed to paper background information about her/him, and read aloud an overall description of the study (without revealing its purpose). Next, participants were trained on using the phone and browser. Training was incremental, starting from simple tasks (e.g., opening the application menu) and ending at two full tasks (e.g., looking from whatis.com at what "ITV" means).

After the training, the trial started. The search task was read aloud to the participant, with the associated bookmark number (e.g., "Choose bookmark number 4"). Some situations involved doing "mobility tasks" related to that location simultaneously with the Web search task (see Figure 1, right column); these were instructed when not implied by the situation. Some tasks were done while moving (route was provided if the participant did not know it) and others while standing or sitting. When moving, the participant led the way and the experimenter shadowed few steps behind without disturbing or helping the participant. After accomplishing the task, the participant's answer was recorded by the experimenter. New instructions were then provided.

Each task was performed in one of the Instructed Time Pressure (ITP) conditions: (1) in the *hurry* condition, the instruction was to "Do as many tasks as quickly as possible." (2) In the *baseline* condition, it was to be done within a given (4 minutes) or implicit timeframe (e.g., "You can do the task until we arrive to the Sörnäinen metro stop"). The timeframe was sufficient to perform the task, but if exceeded, the experimenter stopped the task and instructed the next task. (3) In the *waiting* condition, the participants waited for something, and were told they had plenty of time to carry out the one and only task: "We'll be waiting for a call from my colleague, you have plenty of time."

Altogether, one trial lasted about 1½ hours.

Apparatus

The Web search tasks were performed on a Nokia 6600 mobile phone running a mobile Web browser (Opera).

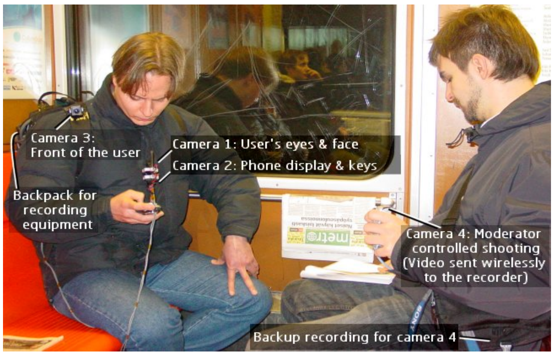


Figure 2. Configuration of recording equipment.

Figure 3. Output video data integrated on-the-fly.

We aimed at making the equipment setup as unobtrusive (for the user and other people) as possible. 30 g (Watek WAT 230A) minicams were used for video recording. One minicam was attached to the test phone, capturing the phone display and keyboard. The device was also equipped with a second camera head that was focused up towards the user's eyes. A third camera was attached to the backpack shoulder strap, facing forward, in order to record the field of vision ahead. Finally, the experimenter's minicam, hidden in a phone shell, captured the overall environment. This video stream was sent wirelessly to a receiver in the participant's backpack. Since we knew that wireless video is susceptible to distractions, we backed up this view onto a tape carried by the experimenter. (See Figure 2.)

The participant carried most of the equipment in a backpack. It contained a microphone, a video camcorder, batteries, a wireless link receiver, and a 4-channel quad processor for building up one video from the four video streams. (See Figure 3, Video Figure, and, for a system diagram, [29].)

Coding

The coders held a calibration meeting where they coded a part of a tape together to agree on and elaborate the coding scheme. Several items were dropped and others simplified to reach a high level of consensus. The revised, final version included measurements that could be done by recognizing or counting specified circumstances (given in the scheme) from a paused video image. Each of the five coders independently coded the tape by watching and pausing the playback after each event and coding it to a data sheet. The final scheme was as follows:

- **Time stamp**: Time for the entry (accuracy of one second)
- **Task number**: 1–25
- **Location**: Café / Metro platform / … (See Figure 1)
- **Instruction on Time Pressure (ITP)**: Hurry / Wait / Normal
- **User's movement**: Walks / Decelerates walk / Stands / Sits
- **Focus of user's attention**: Phone / Environment
- **Interaction**: (User) Starts operating the phone / Stops it
- **Status of loading**: Loading / Page scrollable with only text / All content loaded
- **Crowdedness**: (1) No people around / (2) Some people around (not moving) / (3) Some people around (moving) / (4) Many people moving close, crowded.

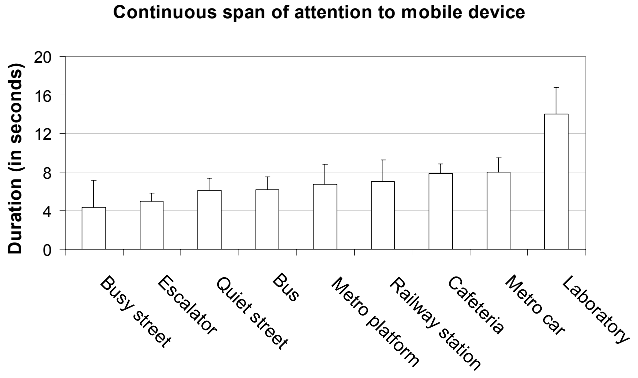


Figure 4. Duration of continuous attention to the mobile device during page loading. Error bars denote 95 % CIs.

RESULTS

(Please see the accompanying Video Figure for examples of data.) An α-level of .05 is utilized throughout the report. Because age differences are not the focus of the present paper, age group differences are reported elsewhere [30].

Page Loadings

Our focus was on the events occurring during the 1894 page loadings, several of which could take place during one task. The response time from link activation to the appearance of the first part of the page was 16.2 seconds in average (mode 7 *s*, median 11 *s*, SD 16 *s*), somewhat biased by long loading times for certain sites. Most variation was caused by different page sizes, variable load, and variance in availability of GPRS connection. This variation is important, as the participants had no possibility of predicting the loading times and adjusting attentional strategies accordingly.

Attentional Resources in Different Mobility Situations

We first examined for how long attention stayed on the mobile device once the page loading had started. Attention shifted away in only 35 % of page loadings in the laboratory, and in the mobile situations from 46 % (Railway station) to 70 % (Metro platform) and 80 % (Long quiet street). A one-way Analysis of Variance (ANOVA) revealed a significant effect of Situation on the *duration of continuous attention* to the mobile device (after the page loading started and before it ended, excluding cases where attention did not shift away from the phone during the page loading), *F*[8,1039]=7.5. A closer examination showed that the laboratory and the café were between 8 and 16 *s*, whereas the escalator and the busy street both fell below 6 *s*. The difference in the duration of continuous attention between the two extremes—the busy street and the laboratory—was over three-fold. (Consult Figure 4 for means and 95 % confidence intervals, CIs.)

Second, we examined how frequently attention switched away from the phone once the loading started. The number of *attention-switches away* from the mobile device during page loading was approximately 8 in the busy street but less than 1 in the laboratory, again a substantial difference. The omnibus *F*-test for the effect of Situation was significant, *F*[8,1871]=22.1. Consult Figure 5 for means and CIs. Interestingly, our hypothesis on the relative order of the escalator and the café gained no support. The café involved more switches than the escalator, perhaps due to (1) turntaking-capture during social interaction in the café and (2) attentional calibration on the escalator (see next subsection).

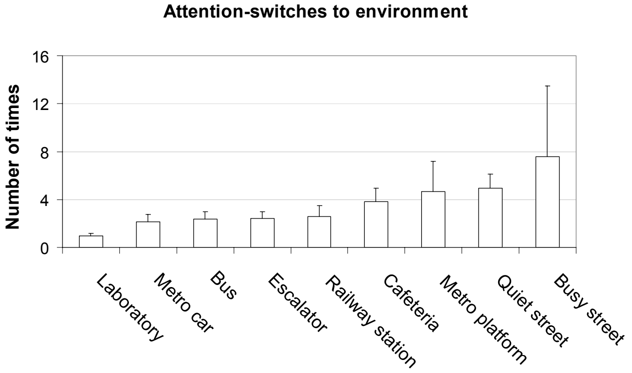


Figure 5. Number of attention-switches away from the mobile device during page loading. Error bars denote 95 % CIs.

Third, *switch-back durations* (the time spent on attending the environment before switching *back* to the mobile device) show only minute differences in how long the environment was attended before switching back to the search task. The omnibus *F*-test on the effect of Situation was not significant, *F*[8,1039]=0.78. In the first of three groups—the laboratory, the metro platform, and the railway station—switch-back duration was in the range of 7 to 8 *s*. The second group— the bus and the café—fell to the range 6–7 *s*, and the third group—the escalator, the quiet street, the metro car, and the busy street—to 4–6 *s*. The difference between the extremes, the escalator (*M*=4.77 *s*) and the laboratory (*M*=7.59 *s*), was borderline-significant according to a Fisher LSD *post hoc* comparison, *p*=0.06, other *p*s>0.1.

Fourth, the predictions of RCF were also evaluated by looking at the *distribution of attention* between the mobile device and the environment. Table 3, which presents distributions in regard to three contextual variables, reveals marked differences between the situations. Regarding the contribution of Situation, the results are quite similar to other measures of attention. Attention dwelled in the environment much more while outdoors than indoors, the difference between the laboratory and the busy street being almost tenfold. The kind of linear relationship between crowdedness and resource-depletion anticipated by RCF was only par-

Situation	%	Level of Crowdedness	%
Laboratory	5	(1) No people	9
Metro car	14	(2) Few still people	6
Bus	16	(3) Few moving people	13
Station	20	(4) Crowded	36
Escalator	20		
Cafeteria	22	**Instruction on Time Pressure**	**%**
Quiet street	23	Baseline	20
Metro platform	24	Hurry	17
Busy street	51	Wait	29

Table 3. Percentage of time spent attending the environment during page loading.

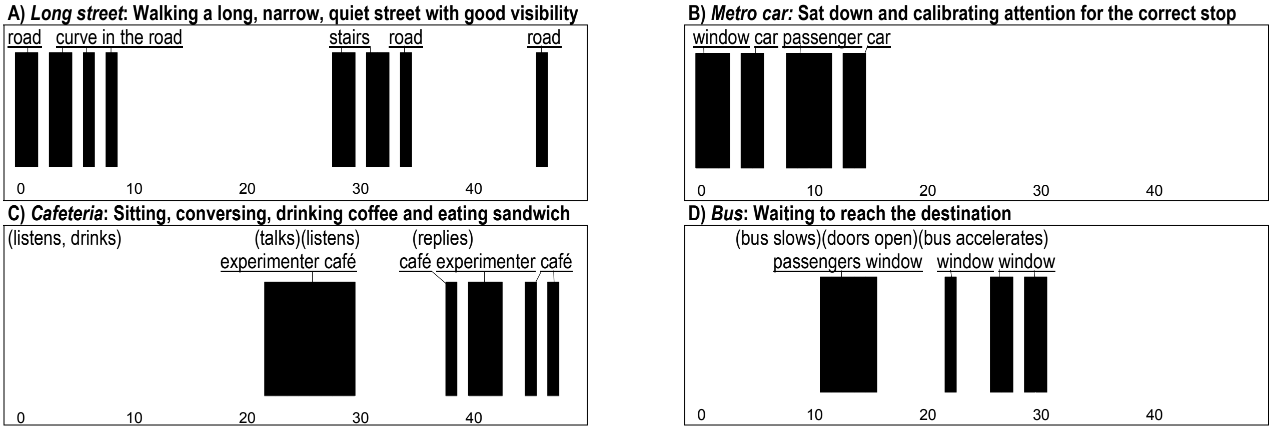


Figure 6. Individual case data to illustrate attentional strategies. Visual gaze patterns are plotted is on a timeline (0-47 *s*), 0 *s* marking the start of page loading. Black bars denote visual attention to the environment, white to the mobile device. All data are particularly long loading times from the same participant (23 years old male student). Observed strategies: Calibrating attention early on: *A)*, *B)*; Brief sampling over long intervals: *A)*, *D)*; Turntaking capture: *C)*.

tially present in the data; the number of people around seemed to matter only when the place was crowded or very crowded. There was only a moderate ($r = -.25$), yet statistically significant, correlation between crowdedness and the duration of continuous attention to the phone. However, we observed that social tasks that involved more turn-taking (e.g., conversation) required more attention. The type of social engagement seems thus more important than the mere number of people present (crowdedness).

Fifth, the results regarding ITPs (Instructed Time Pressure) were close to what was predicted. Table 3 reveals only a negligible decrease of three percentage points due to asking participants to hurry in a HCI task, and an increase of nine points for the waiting ITP. This provides tentative evidence that HCI tasks are lower in the goal hierarchy than mobility tasks. However, although the ITP was associated with the distribution of attention, it did not show correlation with the duration of the *first* continuous span of attention ($r = -.07$). Participants attended to the device equally long in all ITPs before the first switch-away, but after that, the hurrying condition involved a few more and/or longer switch-aways than the other ITPs. The equally long first attention span might indicate calibration (see next subsection) that is needed to establish an attentional strategy for every page loading, and only after this calibration, apparently insensitive to ITPs, resources can be shared more flexibly.

Sixth, we observed, as predicted, that resource depletion forces users to compromise secondary tasks. In particular, participants often had to slow down walking in order to continue interaction with the mobile device. During a page loading, interaction with the device was *stopped* during normal walking on average for 11.12 *s* (95 % CI ± 2.02), but during *slowed* walking for 5.97 *s*, (95 % CI ± 1.97); average page loading times across the two situations being equal. It is natural to reduce the speed of walking in order to reduce the need to frequently sample the environment, which releases resources for other tasks.

Attentional Strategies: Some Qualitative Observations

To fight the negative effects of resource depletion, people have to adopt strategies in attentional switching. RCF does not entail predictive power over these strategies, although it predicts when a need to adopt a strategy emerges. In the following, we report four such strategies observed in qualitative analyses of the video data.

1. Calibrating Attention Early on. As discussed in our earlier work [21], when events in the environment can be predicted accurately on the basis of previous experience, following them does not require as much attention, as they are not surprising. Upon arriving at an environment, we need to calibrate our attention to the predicted events [13]. For example, in Figures 6A and B, one can notice that attention to the environment occurs mostly in the beginning phase of page loading, perhaps because one must estimate how much uninterrupted time for interaction there is left before the next *nodal event* (e.g., arriving at stairs in Figure 6A).

2. Brief Sampling over Long Intervals. After calibrating attention, one can use the environment and its predicted events as pre-programmed reminders to minimize attentional scanning. Predicted (or mentally simulated) nodal events remind of the need to perform an action (e.g., hearing the name of the target station reminds of getting off the metro). Sampling is then reduced to brief sampling over longer intervals, the kind seen in Figures 6A and D. In 6A, the participant is walking on a long street. With brief bursts of perceptual sampling, he can calibrate his attention to the situation where not much is going to happen. Subsequently, during walking, only brief, rapid samples are needed to observe changes that could force a deviation from the plan.

3. Task Finalization. When examining interaction with the device, we observed that participants were often reluctant to attend the environment specifically just before finding the to-be-retrieved information. One explanation is they were unconsciously attempting to minimize cognitive switch costs [37] by decreasing the number of simultaneous tasks.

4. Turntaking Capture. Social interaction requires expressing oneself and orienting to others' responses. Expectations of social behavior guide this interaction [31]. The qualitative difference in the deployment of attention between actively engaging in a discourse and (passively) sampling the *non*-social environment is apparent in our data. For example, compare Figure 6C to the three others: Replying to another person requires temporarily extensive attention and cognitive withdrawal from other tasks. If the task related to social interaction is of higher priority, as it most often is, it easily overrides HCI tasks, leading to long stretches of time with minimal or no attention to the mobile device.

Limitations to the Results

First, an obvious threat to the validity of our conclusions stems from the fact that we were not able to control events taking place in the study locations. Laboratory-like control is simply not possible "in the wild". Some possible countermeasures are discussed in length in [29].

Second, it is possible that our approach *over*estimates attentional resources in mobile situations, and *under*estimates them in the laboratory. The participants took part in a study knowing that they were to be evaluated and measured (evaluation apprehension). This of course affects how the given HCI tasks are prioritized. Under normal circumstances, people would be reluctant to carry out mobile Web search while hurrying through a busy street to catch a bus. Furthermore, the apparatus forced the participant to hold the device at hand all the time, inevitably reducing the threshold for looking at it. Finally, because, in the laboratory, the ability to attend to the device (~14 *s*) was very close to the average page loading time (16 *s*), a ceiling effect for the laboratory condition is likely. The fact that previous research has observed much larger attention spans in office contexts supports this observation (see Discussion).

Third, our measures, confined to visual attention and motor control, are likely to miss the whole picture of compensatory strategies. For example, errors in interaction were not analyzed. In addition, subjective measures of attentional load (e.g., NASA-TLX) could have been used to complement and validate our third-person measures.

Fourth, it is worth pointing out that the deployment of attention is dependent not only on the mobile situation, but on the characteristics of the UI and the main task, and expectations about them. While the reported absolute figures are specific to this particular study, the relative differences among contexts may be more general.

DISCUSSION

During the last two decades, multitasking, multiple user contexts, and related phenomena have become new foci for the HCI research. Thus far, social and cognitive scientists within HCI have concentrated on the topic in their own respective camps [28]. Slowly, however, we are witnessing more evidence that in the shift from desktop interaction toward interaction with ubiquitous and mobile computing neither of the parent sciences suffices on its own. As we have demonstrated, mobile contexts are not only multi-user but also multi-tasking contexts where the social and the cognitive are intimately intertwined. In our view, in mobility, the same social, physical, and artefactual resources and goals that make us desire and enjoy mobility also cause resource-depletion from the point of view of human-computer interaction. This creates a unique challenge where not only attention is limited, but where the user interfaces and computers are as well. Therefore, we believe that mobile HCI makes a truly compelling case for us HCI researchers to build bridges between the two detached disciplines, to "hook" mobile cognition "to the world" [14].

In this paper, we analyzed how attentional limitations emerge due to active participation in the mobile context. The Resource Competition Framework suggests how mobility tasks, often social by nature (e.g., taking care of personal space), reserve the limited cognitive resources and lead to a breakdown in the capacity to interact with mobile devices. The process that decides upon the sharing of attention involves a complex interplay of higher-level mental structures, social expectations for behaving in a certain way, and the psychosocial need to participate in the world.

Our results demonstrate that attentional limitations are very real and seriously constrain mobile interaction. The data conveys the impulsive, fragmented, and drastically short-term nature of attention in mobile interaction. Continuous attention to the mobile device fragmented and broke down to bursts of just 4 to 8 seconds, and attention to the mobile device had to be interrupted by glancing the environment up to 8 times during a page loading! Resources had to be taken from concurrent mobility tasks of secondary importance, for example the control of walking. However, the participants were not giving in to the break down of interaction. We also made observations on rather complex, adaptive strategies to compensate for resource-depletion.

The cost of mobility becomes indisputable when comparing it to the single-user, non-social laboratory condition. The difference is striking, as revealed by the almost eight-fold differences between the lab and the street conditions. Others' recent results support the claim that attention in the office is much less fragmented. For example, looking at office multitasking, Czerwinski *et al.* [7] reported only 0.7 interruptions during a task lasting 53 minutes on average. However, their results, relying on self-reporting and self-generated categories (of e.g. task, switch, and interruption), may underestimate the frequency. Indeed, González and Mark's [11] observation study revealed that information workers spent, on average, three minutes working on one event before switching to another. However, at a higher level of working spheres (set of interrelated tasks), average in-task duration was 11.5 min. They conclude: "work can be described more specifically not as multi-task processing, but as requiring attentional resources to constantly change between different events, tools and working spheres". Going mobile takes this "multi-task processing" to an extreme where interaction breaks down to bursts of just few seconds. These observations are strong evidence for the importance of testing and experimenting in the field (cf. [15]). However, in comparing the office to mobility, one must remember that the two are qualitatively different (e.g., [2,3]). Moreover, micro-level attention studies similar to ours are yet to be done in the office.

Our work also puts forwards a fresh look at the problem of interruptions (i.e., temporary disruptions in the fluent cognitive processing of a task). Whereas researchers in desktop applications and ubiquitous computing have recently understood the importance of the phenomenon, it has been examined from a very performance-oriented viewpoint. The internal constituents of interruptions and task-switches are practically unknown to the field, although, as noted in almost all field studies of interruptions, half or most interruptions are self-initiated [7,11]. RCF stresses that interruptions emerge from delicate balancing of, on the one hand, the limited, hierarchically organized psychological resources and, on the other hand, the requirements of the tasks at hand. Therefore, in addition to social, environmental, and cognitive factors, RCF is concerned with conative (motivational, intentional) mental entities, particularly how they are manifested in the self-deployment of attention between tasks. Our Instructed Time Pressure manipulation results brings some evidence that managing the mobile context is very real and psychosocially important—mobility tasks were observed to override a HCI task. Indeed, the results suggest that participants were not able to withhold managing mobile contexts when asked to do so (by asking them to hurry with the interaction task). In contrast, they *were* able to push the HCI task *down* in the hierarchy of goals and allocate even more attention to the environment when told that they had plenty of time. Resource competition is thus an important factor constituting interruption and task management in mobility and mobile HCI.

We believe that RCF is best suited for examining situations where resource-depletion might emerge. Although our focus has been on attentional resources, and particularly self-deployment of visual gaze, the framework opens a much wider perspective to mobile cognition. In addition to the previously mentioned, there are three research areas of particular interest. First, assessing in detail the "mental workload" involved in managing typical use situations of a device would be beneficial in designing "minimal attention" interfaces for mobile devices [16,25,32]. Second, designing context-aware ubiquitous computers [36] and attentive user interfaces [35] requires an understanding of the relation of mobile use situations to human capacities [20]. Third, RCF can be helpful in predicting "context-triggered" cognitive limitations, which could prove fruitful in the design of multimodal and -sensor interfaces. In general, such systems have the potential of providing much better support for users' context-driven allocation of attention [24]. We envision that our results and the framework are an important step towards developing less attention demanding and interaction techniques for mobile people. Our recent interest has been to utilize the results to design the timing of tactile feedback for page loading in mobile browsers [30].

In this paper, we have also introduced a novel method for studying attention and interaction in mobility. As it now stands, the method is best characterized as a semi-naturalistic field study, but our ambition is to elaborate it toward to what we call a *mobile quasi-experiment* [5], which, through more careful treatment of complex experimental design and of nuisance variables, would allow for making stronger causal inferences. Several lessons on how to develop the methodology were learned. Future work should aim for better external calibration and more rigorous assessment of the reliability of the utilized measures, as well as better documentation of related environmental contexts. Moreover, it should explore a broader range of realistic tasks, including ones that require more continuous attention and interactive participation. In employing the method, we could advance the understanding of mobile cognition by carrying out (1) comparisons between different UI designs in mobile settings—for example, by examining disruption to attention, navigation, conversation, collisions etc., (2) analyses of other, finer-grained response patterns, such as users' refusal to use devices in cognitively or socially taxing mobile situations, and (3) studies of individual differences, such as different attentional strategies in mobility.

ACKNOWLEDGEMENTS

We thank Jaakko Aspara, Tuulia Haikarainen, Barbara Hammond, Tero Jantunen, Harri Lehmuskallio, Miikka Miettinen, Martti Mäntylä, Tuomo Nyyssönen, Petri Piippo, Antti Salovaara, and the six CHI reviewers for generous assistance. The Nokia Multimedia Business Group and the Academy of Finland supported this work financially.

REFERENCES

1. Baddeley, A. D. and Hitch, G. J. Working memory. In G. Bower (Ed.), *The psychology of learning and motivation, Vol. VIII*. Academic Press, NY, 1974, 47-89.
2. Bannon, L., Cypher, A., Greenspan, S. and Monty, M.L. Evaluation and analysis of users' activity organization. In *Proc. CHI'83*, ACM Press (1983), 54-57.
3. Bellotti, V. and Bly, S.A. Walking away from the desktop computer: Distributed collaboration and mobility in a product design team. In *Proc. of CSCW'96*, ACM Press (1996), 279-286.
4. Card, S. Moran, T. and Newell, A. *The psychology of human-computer interaction*. Lawrence Erlbaum, 1983.
5. Cook, T.D. and Campbell, D. *Quasi-experimentation: Design and analysis issues for field settings*. Rand McNally, 1979.
6. Cowan, N. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 1 (2001), 87-185.

7. Czerwinski, M., Horvitz, E. and Wilhite, S. A diary study of task switching and interruptions. In *Proc. CHI'04*, ACM Press (2004), 175-182.

8. Deci, E.L. and Ryan, R.M. *Intrinsic motivation and self-determination in human behavior.* Plenum Press, New York, 1985.

9. Dourish, P. *Where the Action Is: The Foundations of Embodied Interaction.* MIT Press, Cambridge, 2001.

10. Ericsson, K. A. and Kintsch, W. Long-term working memory. *Psychological Review, 102* (1995), 211-245.

11. González, V.M. and Mark, G. "Constant, constant, multi-tasking craziness": managing multiple working spheres. In *Proc. CHI'04*, ACM Press (2004), 113-120.

12. Goodwin, C. and Duranti, A. Rethinking context: an introduction. In A. Duranti & C. Goodwin (Eds.), *Rethinking Context: Language as an Interactive Phenomenon*. Cambridge Univ. Press, Cambridge, UK, 1992, 1-42.

13. Harris, J. and Wilkins, A.J. Remembering to do things: A theoretical framework and illustrative experiment. *Human Learning, 1* (1982), 1-14.

14. Hutchins, E. *Cognition in the Wild*. MIT Press, Cronbridge, MA, 1995.

15. Kjeldskov J., Skov M. B., Als B. S. and Høegh R. T. Is it worth the hassle? Exploring the added value of evaluating the usability of context-aware mobile systems in the field . In *Proc. Mobile HCI 2004*, 61-73.

16. Kristoffersen, S. and Ljunberg, F. Making place to make it work: Empirical exploration of HCI for mobile CSCW. In *Proc. GROUP'99*, ACM Press (1999).

17. Kulomäki J. and Summala H. *Multitasking and Attentional Demands in Pedestrian Behaviour - A Review.* Dep. of Psychology, Univ. of Helsinki. 30.6.2004.

18. Lumsden, J. and Brewster, S. A paradigm shift: Alternative interaction techniques for use with mobile & wearable devices. In *Proc. CASCON'03*, 2003.

19. Navon, D. and Gopher, D. On the economy of the human-processing system. *Psychological Review, 86* (1979), 214-255.

20. Oulasvirta, A. and Salovaara, A. A cognitive meta-analysis of design approaches to interruptions in intelligent environments. *Ext. Abstracts CHI'04*, ACM Press (2004), 1155-1158.

21. Oulasvirta, A. and Tamminen, S. Temporal tensions in human-computer interaction. *CHI'04 Workshop on Temporal Aspects of Work*, 2004.

22. Oulasvirta, A. Task-processing demands and memory in web interaction: A levels-of-processing approach. *Interacting with Computers, 16*, 2 (2004), 217-241.

23. Oulasvirta, A. *Human-computer interaction in mobile context: a cognitive resources perspective*. Licentiate Thesis, Cognitive Science, Department of Psychology, University of Helsinki, 2004. Available from http://ethesis.helsinki.fi .

24. Oviatt, S., Coulston, R. and Lunsford, R. When do we interact multimodally? Cognitive load and multimodal communication patterns. In *Proc. ICMI 2004*, 129-136.

25. Pascoe, J., Ryan, N. and Morse, D. Using while moving: HCI issues in fieldwork environments. *ACM TOCHI*, 7, 3 (2000), 417-137.

26. Pashler, H. Dual-task interference and elementary mental mechanisms. In D. Meyer & S. Kornblum (Eds.), *Attention and performance XIV.* MIT Press, MA, 1993.

27. Perry, M., O'Hara, K., Sellen, A., Brown, B. and Harper, R. Dealing with mobility: understanding access any anytime, anywhere. *ACM TOCHI, 8*, 4 (2001), 323-347.

28. Rogers, Y. New Theoretical approaches for Human-Computer Interaction. *Annual Review of Information, Science and Technology, 38* (2004), 87-143.

29. Roto, V., Oulasvirta, A., Haikarainen, T., Kuorelahti, J., Lehmuskallio, H. and Nyyssonen, T. Examining mobile phone use in the wild with quasi-experimentation. *HIIT Technical Report 2004-1*, 2004. www.hiit.fi/publications

30. Roto, V. and Oulasvirta, A. Need for non-visual feedback with long response times in Mobile HCI. A full paper submitted to *WWW 2005*, May 10--14, Japan.

31. Sacks, H., Schegeloff, E.A. and Jefferson, G. A simplest systematics for the organization of turn-taking conversation. *Language, 50* (1974), 696-735.

32. Sawhney, N. and Schmandt, C. Nomadic radio: speech & audio interaction for contextual messaging in nomadic environments.*ACM TOCHI 7*,3 (2000), 353-383.

33. Suchman, L.A. *Plans and situated actions: The problem of human-machine communications*. Cambridge University Press, Cambridge, UK, 1987.

34. Tamminen, S., Oulasvirta, A., Toiskallio, K. and Kankainen, A. Understanding mobile contexts. *Personal and Ubiquitous Computing, 8*, 2 (2004), 135-143.

35. Vertegaal, R. Attentive user interfaces. *Communications of the ACM, 46*, 3 (2003), 31-33.

36. Weiser, M. The computer for the twenty-first century. *Scientific American*, 1991, 94-104.

37. Welford, A.T. The psychological refractory period and the timing of high speed performance-A review and a theory. *British Journal of Psychology, 43* (1952), 2-19.

38. Wickens, C.D. Processing resources in attention. In R. Parasuraman and R. Davies (Eds.), *Varieties of Attention*. Academic Press, New York, 1984.

39. Young, M.S. and Stanton, N.S. Malleable attentional resources theory: a new explanation for the effects of mental underload on performance. *Human Factors, 44*, 3 (2002), 365-375.