

The Impact of Indoor/Outdoor Context on Smartphone Interaction During Walking

Jakub Krukar
Institute for Geoinformatics
University of Muenster
Heisenbergstr. 2
Muenster, Germany
krukar@uni-muenster.de

Antonia van Eek
Institute for Geoinformatics
University of Muenster
Heisenbergstr. 2
Muenster, Germany
antonia.vaneek@uni-muenster.de

Abstract

It is unclear how users' interaction patterns with personal technology change, as they move between indoor and outdoor spaces. Understanding the impact of indoor/outdoor context could help to improve adaptive user interfaces of location-based services. We present a field experiment, in which participants were asked to complete a cognitive task appearing on a smartphone, while walking subsequent indoor and outdoor route segments. Results demonstrate that the average time to complete the task was significantly longer, when the same task was encountered outdoors, compared to indoors. The recorded eye tracking data show that the rate at which participants were glimpsing the smartphone (anticipating the task to appear on the screen) was decreasing steadily over time, but increased again after re-entering an already known building. This indicates that the indoor/outdoor context has a significant impact on at least two distinct aspects of interacting with personal technology during walking: the user's ability to quickly respond to a simple task requiring thinking, and the user's spontaneous willingness to visually control the display. Most importantly, it is shown that the influence of the indoor/outdoor context on these two aspects manifests itself with two distinct patterns. The findings are discussed with regard to different attentional requirements of indoor and outdoor spaces during walking.

Keywords: user context, smartphone interaction, eye movement

1 Introduction

Location-based services (LBS) and location-aware software applications offer the promise of adjusting the content and style of the interaction to the current geographical context of the user. However, integration of indoor and outdoor contexts is known to be a challenge for the future development of LBS (Huang *et al.*, 2018). This paper considers the problem of interacting with smartphones while walking: an important aspect of the practical use of LBS technologies.

It is unclear how users' interaction patterns with smartphones change, as they move between indoor and outdoor spaces. Since LBS applications are often used while walking, their use is likely associated with poorer cognitive performance and poorer situation awareness (Lin and Huang, 2017). However, attentional demands of indoor and outdoor spaces are different (Kray *et al.*, 2013). Indoor navigation, for instance, is associated with a more often need for making navigational decisions and the similarity of the surrounding environment might make these decisions more difficult, compared to the outdoor context (Hölscher in: Krukar, Hölscher and Conroy Dalton, 2017). It thus seems, that, keeping all other things equal, using a smartphone while navigating indoors might be associated with a different cognitive demand than using it while navigating outdoors.

However, it is difficult to study the exclusive influence of the indoor/outdoor context on software interactions because the way location-based services work is different indoors compared to outdoors. For example, the interface typical to outdoor navigation systems is not optimal for indoor navigation (Huang and Gartner, 2009; Puikkonen *et al.*, 2009)

and thus successful indoor navigation systems deviate from their outdoor counterparts' standards. Any change in the human-smartphone interaction indoors, compared to outdoors, might therefore come from a different look of the application, and not from the user's context alone.

In order to tackle this issue, we study the influence of the indoor/outdoor context using a simple cognitive task that remains unchanged indoors and outdoors. It can be therefore assumed that any change in task performance is the result of the user's context while performing the task, and not the changing difficulty of the task itself.

We hypothesised that, as participants move from indoor to outdoor spaces, their task performance would be better, i.e., the average speed at which they answer questions would decrease. Since the reason for that change might lie in the distribution of attention between the task and the surrounding environment while walking, we employ mobile eye-tracking (Kiefer *et al.*, 2017) for recording participants' gaze behaviour.

2 Method

2.1 Participants

Fifteen participants (7 male, 8 female; aged 19 - 34, $M = 26.4$) were recruited through a university's newsletter. Data of 2 could not be analysed due to technical issues. Each person was compensated with 10 euros.

2.2 Materials

Participants followed a route that consisted of 5 segments: Building 1a, Street Section 1, Building 2, Street Section 2, Building 1b. Sections named Building 1a and 1b were two different routes in the same building, running between its main entrance and the laboratory room, at which the experiment started and ended. Each participant followed exactly the same route. Routes inside the buildings were taking participants through smaller corridors that were never crowded by other building users at the time of the experiment. Routes outside the buildings were on sidewalks and pedestrian areas, with two instances of a traffic light crossing.

While walking, participants were required to respond to questions appearing on a smartphone carried by them in their hand. The application displayed a random question from its fixed set of 94 questions, together with 3 possible answers. The questions stayed visible on the display until the participant provided an answer by tapping one of the solutions. After providing the response, the question disappeared. A new question would appear after a random period of time between 10 and 40 s. The screen remained blank otherwise. Sample questions (answers) were: "What is the colour of the sky? (blue, yellow, red)", "585 - 263 = ... (322, 741, 243)", "What is the capital of Portugal? (Lisbon, Cairo, Tel Aviv)". Each question was presented to each participant maximum once. The questions were purposefully designed as solvable, since our main interest was in the reaction times, not in the accuracy of the answers.

2.3 Experimental Design and Procedure

We employed a within-subject experimental design. Each participant was met at the laboratory room in Building 1. After signing an informed consent form, they were asked to wear a mobile eye-tracker, a backpack with a laptop running the eye-tracking software, a sun-protective hat, and underwent the calibration procedure. Next, they were given the smartphone and explained that their task is to answer questions appearing on its screen as quickly and as accurately as possible. They did not know when the next question may appear exactly, except that it will appear between 10 and 40 seconds after the previous one. Each participant would see each question maximum once.

Participants were then asked to follow the experimenter. They did not know what route they will be taking, except that it should take around 30 minutes. They were asked to always prioritise their own safety and only cross streets with the experimenter.

Since the order of questions was randomised, it was possible that different questions would appear in different spaces for each participant. We utilise the statistical method of mixed-effect models in order to control for the effect of individual question. This makes it possible to analyse the isolated effect of the indoor/outdoor context on participants' performance even if questions differed in their difficulty.

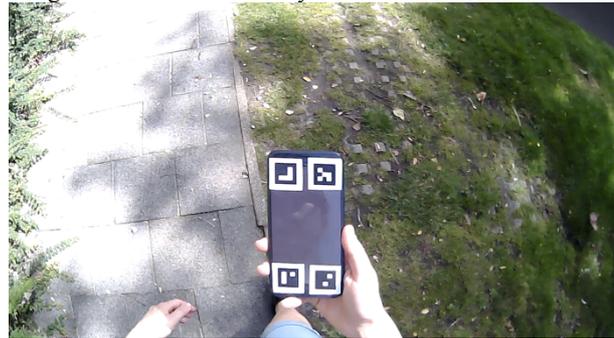
2.4 Apparatus

We used a Pupil Labs binocular mobile eye-tracker with the reported accuracy of 0.6 degree of visual angle (Kassner, Patera and Bulling, 2014). Gaze was recorded at the resolution

of 120 Hz with activated function for autocorrecting changes in lighting. The world camera was pointed downwards, in order to capture a smartphone held at the waist height within its field of view.

The task was presented using a custom-built Android application on a Google Pixel smartphone with a 5-inch screen. Four square markers were attached to the edges of the display, allowing the Pupil Labs software to automatically detect gaze occurring on the smartphone (Figure 1).

Figure 1: A view from the eye-tracker's world camera.



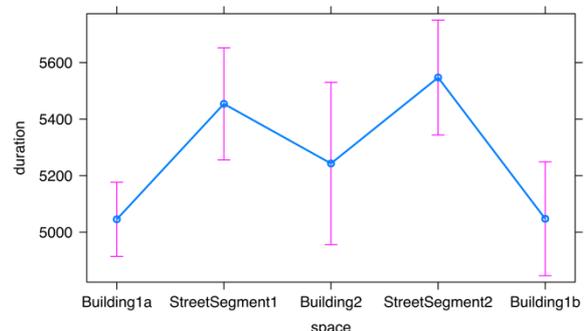
3 Results

3.1 Task Response Time

We were primarily interested in Task Response Time, i.e., the time it took participants to provide an answer to each question, from the moment it appeared on the screen. For this analysis we only consider correct responses, which is the standard approach in response time analysis. Average overall Task Response Time equalled $M = 4645$ ms, $SD = 3270$ (4556 ms in Building 1a, 4641 ms in Street Segment 1, 4442 ms in Building 2, 4937 ms in Street Segment 2, and 4080 ms in Building 1b).

We used linear mixed-effect models with an inverse gaussian link to analyse these differences. This approach considers the fact that each participant provided multiple responses, but to a limited subset of possible questions. The result is the model's predicted response time to an average question, by an average participant, in each space (Fig. 2).

Figure 2: Predicted Response Times for each space.



In order to directly test our hypothesis, we analysed the significance of the differences between indoor vs. outdoor

spaces. Table 1 presents the results of a linear mixed-effect model, where Building 1a, 2, and 1b are jointly considered as "indoor spaces" and Street Segments 1 and 2 are jointly considered as "outdoor spaces". As visible, after accounting for large variance coming from the random effects of participant ID and question ID (i.e., for a large variance in performance between participants, and large variance in the difficulty of individual tasks), the difference between two types of spaces is significant: Answering to a question while being outdoors took significantly longer (around 391 ms longer) than answering to the same question indoors. This is contrary to our initial hypothesis.

Table 1: Linear mixed-effect model of Task Response Time.

Task Response Time			
Predictors	Estimates	CI	p
(Intercept)	5317.82	5167.64 – 5468.00	<0.001
outdoor (compared to indoor)	391.44	91.09 – 691.80	0.011
Random Effects			
σ^2	0.00		
τ_{00} question.ID	1327241.00		
τ_{00} participant.id	280819.94		
ICC question.ID	0.83		
ICC participant.id	0.17		

3.2 Eye Movement Controlling the Display

In order to explain the differences in Task Response Times, we analysed gaze data from the mobile eye tracker. We hypothesized that longer Task Response Times are caused by less frequent glimpsing at the display: a participant who is checking the blank smartphone display less often, would be more delayed in providing the answers.

Since participants were asked to move between spaces with different lighting conditions, there was a large data loss in the eye tracking recordings that may be systematically related to each space. Following the recommendation of (Evans *et al.*, 2012) we only analysed aggregated eye tracking data. We developed an aggregated dependent variable by dividing the number of fixations on the surface of the smartphone (in each space, by each participant) by the number of all fixations detected (in the given space, for the given participant). This resulted in a measure that is unbiased by the possibly uneven data loss across indoor and outdoor spaces. The resulting variable (called *Controlling Ratio* below) describes how often each participant glimpsed at the display of the smartphone, checking whether a new task has appeared on the screen. A higher value indicates that the participant was glimpsing the display more often. Our hypothesis therefore stated that participants should glimpse at the device more often in the indoor spaces, which had led to faster Task Response Times.

The analysis omits fixations that occurred on the smartphone display when a question was visible on the screen, since these fixations were used to read the task, and not to control the display for the presence of a new question.

The eye tracker detected, on average, $M = 1275$ ($SD = 347$) fixations per participant, within which $M = 220$ ($SD = 121$) fixations on average happened on the surface of the display. The average Controlling Ratio equalled 0.17 (0.26 in Building 1a, 0.18 in Street Segment 1, 0.14 in Building 2, 0.11 in Street Segment 2, and 0.17 in Building 1b).

We used linear mixed-effect models to investigate the significance of these differences (Table 2). As can be seen on Fig. 3, participants were steadily decreasing the rate at which they were glimpsing at the display, until they have moved back to Building 1 again.

Figure 3: Predicted Controlling Ratio in different spaces.

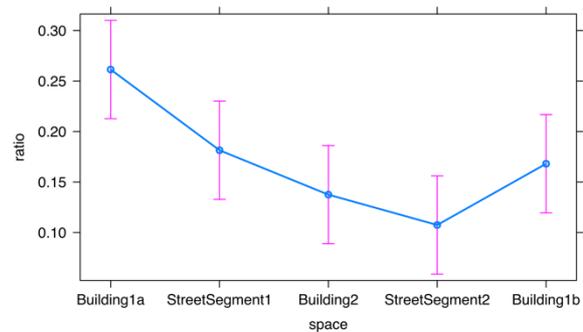


Table 2: Linear mixed-effect model of Controlling Ratio. All estimates and significance values show the difference compared to the base-line (Building 1a).

Controlling Ratio			
Predictors	Estimates	CI	p
(Intercept)	0.26	.21 – .31	<0.001
Street Segment 1	-0.08	-.13 – -.03	0.002
Building 2	-0.12	-.17 – -.07	<0.001
Street Segment 2	-0.15	-.20 – -.11	<0.001
Building 1b	-0.09	-.14 – -.04	<0.001
Random Effects			
σ^2		0.00	
τ_{00} participant.id		0.00	
ICC participant.id		0.47	

3.3 Task Accuracy

Overall Task Accuracy equalled 95% (91% in Building 1a, 96% in Street Section 1, 98% in Building 2, 95% in Street Segment 2, and 94% in Building 1b). We used linear mixed-effect models with a binomial link to measure the impact of indoor/outdoor context on task accuracy (Table 3). There was a statistically significant difference between Building 1a and Street Section 1, but differences between other pairs were not significant.

Table 3: Linear mixed-effect model of task accuracy. All estimates and significance values show the difference compared to the base-line (Building 1a).

Predictors	Task Accuracy		
	Odds Ratios	CI	p
(Intercept)	725.83	57.91 – 9098.1	<0.001
Street Section 1	6.94	1.17 – 41.02	0.033
Building 2	8.37	0.92 – 76.21	0.059
Street Section 2	4.40	0.81 – 23.95	0.087
Building 1b	3.17	0.30 – 33.54	0.337
Random Effects			
σ^2	3.29		
τ_{00} interaction window ID	33.72		
τ_{00} participant.id	0.15		
ICC interaction window ID	0.91		
ICC participant.id	0.00		

4 Discussion

Participants were slower when responding to questions seen outdoors, compared to a situation when they were encountered indoors (after controlling for the random effect of the individual question). This difference, however, was not fully explained by the eye tracking data. Contrary to our second hypothesis, participants did not control the display less frequently outdoors, than indoors. Instead, they steadily decreased the frequency of controlling the display over time. This trend was reversed when they re-entered a building already known from the first part of their journey.

These results demonstrate that two cognitive processes (thinking about the solution, and visually paying attention to the smartphone's display), follow two different patterns when smartphone users move between indoor and outdoor spaces. One potential explanation is that outdoor spaces present higher mental workload during walking, possibly because of the higher number of surrounding stimuli, higher danger of interacting with car traffic, and larger viewing fields extending from the participant in all directions. In contrast, indoor spaces present a restricted number of visual stimuli

while walking: one only needs to pay attention to the corridor ahead, and the size of one's viewing field is limited by walls. Active navigation was not required in our experiment (since participants simply followed the experimenter): this might explain the significant difference between our results and the assumptions from the existing literature. In this scenario, it seems that as participants became more confident with the task, they tended to shift more attention to the environment, at the expense of controlling the display.

The biggest limitation of the current work is a small sample size and a restricted number of the studied environments. Further studies with increased sample sizes could help clarify the reasons behind the observed effects.

5 Conclusion

The results of this study have potential implications for the design of location-based services used across indoor and outdoor spaces. It can be concluded, that the same simple mental task requires more time to perform when encountered outdoors. Technology designers could thus simplify the type of input required outdoors in their location-aware software. The eye movement data demonstrates, that software designers can also expect a spike in the attentional engagement with the display when users re-enter already known spaces. Location-aware applications that preserve the history of recent locations could use this opportunity to present users with more information, than it would be otherwise manageable.

Our future work will extend this approach to a larger number of spaces, and more diverse cognitive tasks that could better resemble activities that users of location-based services are required to perform on their smartphones.

6 Acknowledgments

This research was supported by the ERC StRG Grant Agreement No 637645.

Jakub Krukar kindly acknowledges the support of AGILE's early stage career development bursary.

References

- Evans, K. M., Jacobs, R. A., Tarduno, J. A. and Pelz, J. B. (2012) 'Collecting and Analyzing Eye-tracking Data in Outdoor Environments', *Journal of Eye Movement Research*, 5(2), pp. 1–19.
- Huang, H. and Gartner, G. (2009) 'A Survey of Mobile Indoor Navigation Systems', in *Cartography in Central and Eastern Europe*. Springer, pp. 305–319. doi: 10.1007/978-3-642-03294-3_20.
- Huang, H., Gartner, G., Krisp, J. M., Raubal, M. and Van de Weghe, N. (2018) 'Location based services: ongoing evolution and research agenda', *Journal of Location Based Services*. Taylor & Francis, 00(00), pp. 1–31. doi: 10.1080/17489725.2018.1508763.
- Kassner, M., Patera, W. and Bulling, A. (2014) 'Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction', *Proceedings of the 2014 ACM*

International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, pp. 1151–1160. doi: 10.1145/2638728.2641695.

Kiefer, P., Giannopoulos, I., Raubal, M. and Duchowski, A. (2017) ‘Eye tracking for spatial research: Cognition, computation, challenges’, *Spatial Cognition & Computation*, 17(1–2), pp. 1–19. doi: 10.1080/13875868.2016.1254634.

Kray, C., Fritze, H., Fechner, T., Schwering, A., Li, R. and Anacta, V. J. (2013) ‘Transitional Spaces: Between Indoor and Outdoor Spaces’, in *Spatial Information Theory*. Scarborough, UK: Springer, pp. 14–32.

Krukar, J., Hölscher, C. and Conroy Dalton, R. (2017) ‘Indoor Wayfinding: Interview with Christoph Hölscher and Ruth Conroy Dalton’, *German Journal of Artificial Intelligence (Künstliche Intelligenz)*, 31(2), pp. 185–191. doi: 10.1007/s13218-016-0483-3.

Lin, M.-I. B. and Huang, Y.-P. (2017) ‘The impact of walking while using a smartphone on pedestrians’ awareness of roadside events’, *Accident Analysis & Prevention*. Pergamon, 101, pp. 87–96. doi: 10.1016/J.AAP.2017.02.005.

Puikkonen, A., Sarjanoja, A.-H., Haveri, M., Huhtala, J. and Häkkilä, J. (2009) ‘Towards designing better maps for indoor navigation’, in *Proceedings of the 8th International Conference on Mobile and Ubiquitous Multimedia - MUM '09*. New York, New York, USA: ACM Press (MUM '09), pp. 1–4. doi: 10.1145/1658550.1658566.