Running: A Flexible Situated Study

Elizaveta Temir
Aisling Ann O'Kane
Paul Marshall
Ann Blandford

University College London
London, United Kingdom
elizaveta.temir.12@ucl.ac.uk
a.okane@cs.ucl.ac.uk
paul.marshall@ucl.ac.uk
a.blandford@ucl.ac.uk

Abstract

This paper describes a situated study of personal informatics applications for running that had to be conducted in a flexible and pragmatic way to adjust for the context of use. A qualitative situated study highlighted important differences in runners' motivations, uncovering markedly different uses and preferences between people who run either for health or for pleasure, but also underscored how the physical nature of the interaction impacted data collection. By adjusting the method to be sensitive to the physical nature of the interaction and the preferences of the participants, a pragmatic situated approach provided insights into how these technologies are actually used.

Author Keywords
Wellbeing; fitness; exercise; running; situated studies; apps; context; mobile; personal informatics.

ACM Classification Keywords
H.5.m. Information interfaces and presentation.

Introduction
The market for mobile interactive technologies to support sport, fitness, and wellbeing, such as wearable technologies, activity tracking tools, and smartphone apps, is growing rapidly. These systems can track physical activity, so their context of use can be mobile for those who choose outdoor activities such as running. The context of use of mobile devices has been shown to significantly affect user experience [25] and there has been a push to evaluate mobile apps in-situ [11]. However, context is a dynamic product of interaction: it emerges from on-going activity and cannot be separated from it [7]. The context of running includes the physical environment, an individual's physical and emotional state, past and present interactions with other apps, experience as a runner, personal goals, and technology used to support running that are all specific to a particular situation. Furthermore, outdoor running occurs in environments that are subject to rapid, unpredictable and

uncontrollable changes and poses many physical and other constraints on the use of technology.

Understanding the details of how running applications are used in these changing contexts would give insight into how well they support runners (cf. [3]). Existing studies of mobile applications have tended to be retrospective, using methods such as interviews [12] or diary studies [2], or sampling snippets of use through experience sampling [18]. However, there are two drawbacks to these methods: lack of coverage of all interactions and lack of detail on the interactions [1]. Moreover, mobility and small screen size make traditional observational methods difficult in dynamic contexts. To overcome these challenges, Brown et al. [1] suggested a novel approach using a combination of wearable video cameras and screen capture to study the details of interaction with a mobile map application. We attempted to replicate this, however we found that we had to significantly adapt it because of the physical activity and preferences of the participants. Using a more flexible and pragmatic situated approach is promising as it revealed parts of the experience with the app whilst exercising that usually remain invisible and helped us to evaluate applications when they are actually used, during physical activity and exercise.

Background

As activity tracking systems become more popular, researchers have increasingly examined how personal informatics systems (PIs) [16] are weaved into people's everyday lives: a perspective characterised as *"lived informatics"* [26]. In-situ research 'in the wild' has focused on deploying novel technology prototypes and studying their effects on user behaviour (e.g. [4–6,17]). Another focus has been on users of commercial

PIs, including the barriers that they experience – such as tracking accuracy [10,28], issues with device aesthetics in different contexts [10], lack of fit with users' self perception, and challenges in using data [15]. Other work has focused on specific groups, such as people who persevere in using activity trackers over long periods [9] and elite athletes who use them to measure performance [27]. Initial attempts to capture the embodied nature of exercise have included the use of a quad-copter jogging companion [20], field studies of a back mounted display for group running [19], and a situated study of physical exercise in the gym [23]. However, most studies are separated from the context of use [13] or rely on retrospective accounts by the participants (e.g. [8,9,14,26]).

Brown et al [1] discuss two limitations of not taking a situated approach. Firstly, lack of coverage: these methods are unlikely to cover important situations of use, as both interviews and diary studies are retrospective and interviewees will remember and prioritise some instances over others, and experience sampling only gets a snapshot of a situation. Secondly, a lack of detail: interviewees filter retrospective accounts based on what they find important and diary studies are limited in how much can be asked of respondents. Instead, they suggest using wearable video cameras to capture interactions with the device and the surrounding context of use, along with screen capture. Brown et al.'s approach inspired us to investigate the use of fitness technology where PIs are actually used 'in the wild': when running outdoors.

A Pragmatic Situated Approach

We approached the question of how running applications were used in-situ through an exploratory
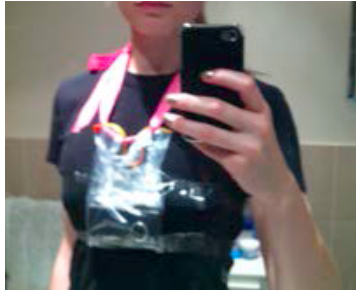
Figure 1: First author wearing camera in a bag on her chest to capture the use of running apps on her mobile phone in-situ

situated qualitative study. Two free versions of the running apps Endomondo and Runtastic were chosen due to their market dominance at the time of the study. To get as close as possible to the context of use, we used a range of situated methods: autoethnography and user studies, using participant observation, contextual interviews, and participant diaries.

Autoethnography was employed by the first author to understand subtle interaction issues [21] and gain empathy for participants [22], as well as to pilot the user methods. Based on initial user pilot tests, screen capture technology was strongly rejected because of the potential to capture personal data, but following Brown et al.'s method [1], she wore a camera in a plastic bag on her chest in order to collect video data on her use of apps, as she tended to hold her mobile phone in her hand (see Figure 1). Although she received plenty of confused looks from passers-by, the data collected from the video was found to be useful. To improve field-testing methodologies, a week-long user pilot study was carried out. Contextual interviews proved to be appropriate, but video capture using the wearable camera setup was rejected as it did not capture phone use when in an arm pouch, the bag worn around the neck bounced when running, taping the camera to clothing left marks on the pilot participant's clothing, and the pilot participant also suggested that any kind of wearable solution (e.g., a GoPro on a head strap) would be disruptive to the run itself.

For the user study, the researcher had three points of contact with participants. The first was a contextual interview, and then a joint run (observation with think-aloud) followed by a post-run contextual interview (all captured by the researcher with a hand-held camera).

Then the participant ran as usual using an app for two weeks and kept a diary (text, voice notes, and/or photos) about their interactions using the same phone. Diary entries were made immediately after the run. The second meeting was after two weeks and started with a pre-run contextual interview about experiences with the second app (half of the participants used Endomundo first, half used Runtastic first), where participants went through their diary entries. Then they ran with the researcher while using the second app (observation with think aloud), followed by a post-run contextual interview (all captured with a hand-held camera). They then ran alone using the second app for another two weeks and kept a diary of their app use. Finally, an exit interview was conducted where participants discussed the experiences recorded in their diary entries for the second app; an open-conversation format was used to probe for comparisons between the apps.

After the male pilot participant, six participants (three males) aged from 23 to 33 years old (M=27) were recruited for the user study through snowball sampling. All were regular runners who already used activity tracking in their routines [24]. Participants ran between one and six times with each app in the two-week period keeping a diary, with an average of three runs each with Runtastic and four runs each with Endomondo. Participants were given £20 for their participation.

Based on issues replicating Brown et al's approach [1], participant observations with video (V), contextual interviews (CI), and the diary study (DS) of the use of both apps were adapted to the physical constraints of this study so that we could gain access to both self-reported issues and specific interactions in the context of outdoor running.
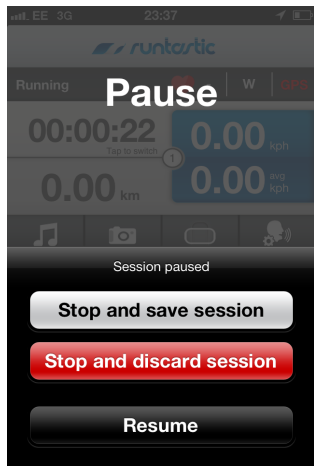
Figure 2: The "Stop and save session" button was too close to the "Stop and discard session" button for all participants with shaky hands after a run

Findings

Expectations Based on Prior Experience and Priorities
Participants had clear expectations of the apps that they were using in this study. If there was a feature they had liked in a previously used app, they were repeatedly annoyed by the lack of it, particularly if it influenced running. The live map of one application was surprisingly not interactive, which was annoying considering the mobile nature of outdoor running: *"The app did frustrate me because it showed me where I was, but it didn't let me move the map around"* – P1, D; *"I basically have to remember where I am going or check the map every two blocks"* – P4, V.

Other than these universal expectations based on prior experience, there were expectations that differed significantly based on motivation to run. We found that people who self-identified as running for personal health (P1, P2, P3) during the contextual interviews wanted free features to motivate them towards a goal, which was often distance: *"I ran two kilometres, and after that it said "to get more voice feedback [pay for pro]" and then it didn't tell me that I was on my third kilometre […] that kind of annoyed me"* - P2, CI.

Motivation for running influenced whether some features were seen as a distraction or a disruption, (cf. [23]). Where some *"don't listen to music when [they] run"* (P6, CI), others thought that not being able to *"even play music without upgrading […] is bullshit"* (P1, CI). Some features could get in the way of the enjoyment of the people who self-identified as running for pleasure in the contextual interviews like P4, P5, and P6, which is in contrast to those wanting motivation when not running for pleasure, but rather for health, like P1, P2, and P3:*"I quite enjoy running

without an app because you just go and get it done. […] lots of times it can interfere, and that's why I've not used it religiously"* – P6, CI; *"Maybe since running isn't so much pleasure for me […] So right now I feel like I need a little bit more motivation from the app, a little bit more encouragement.* – P1, DS.

In summary, for features commonly found on non-fitness apps, expectations were similar across all users. However, there were significant differences between expectations that were associated with distraction or motivation that health runners found encouraging and pleasure runners found annoying and disruptive.

Engagement Influenced by Embodied Nature of the Run
The embodied nature of running with a physical device influenced its use. For instance, two buttons were too close together when one app was paused (Figure 2), making it easy discard data rather than save it, leaving P1 *"a bit terrified"* (DS). Noticeably frustrated during a pause in running during the observation, hobby runner P5 described the situation: *"This is a bit stupid, because stop and save is really close to stop and discard. If you're still a bit high after your run, and a bit shaky, you might hit the wrong button and accidentally delete the whole run."* – P5, V.

Another situational factor that impacted engagement with the apps was the location of interface buttons near the edges of the screen, which did not take into account interaction when the device was attached to an arm. Health runner P3 attempted tapping and buttons numerous times whilst looking increasingly frustrated and reported: *"These buttons the corners are really difficult […] especially when you run with it [pointing to iPhone arm pouch]."* – P3, V.
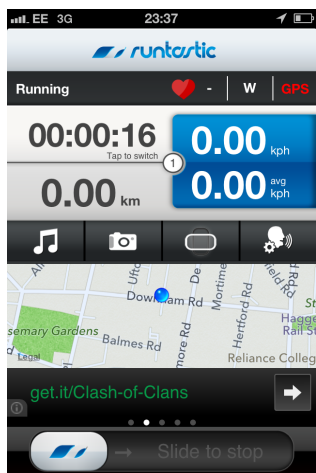
Figure 3: Pleasure runner P5 wanted to get rid of distance as he found it cluttered while health runner P3 wanted a larger representation of distance.
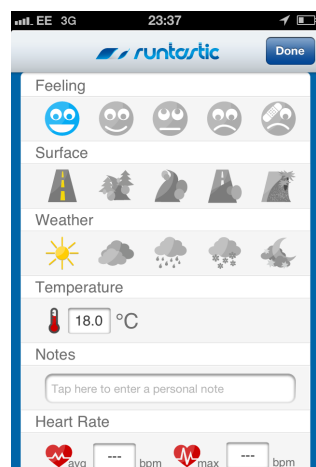

Figure 4: Extra step before saving that was useful for pleasure runners and annoyed health runners

Engagement with the apps when running was also influenced by people's motivation to run. They were trying to use the interface for different purposes, such as pleasure runners like P5 not wanting to see distance in the moment and health runners like P3 prioritising this goal related information and wanting a bigger representation of it (Figure 3) : *"They could probably do away with [...] distance. It's a little bit too cluttered here. I don't like it"* – P5, V; *"When it shows you the distance on screen while you're running. [...] It's quite tiny. I probably would have missed that"* - P3, CI.

Like their expectation for motivational support from the apps, those trying to reach health goals wanted to engage with the apps in a more personal way. When discussing her runs, P3 described *"hate for the robot voice"* and found that it influenced how she felt about reaching her goals: *"Your app is like your running companion [...] the voice was just so – it's so impersonal, because it's not human. You kind of feel like it doesn't really care"* - P3, CI.

The participants who ran for pleasure did not want a 'companion.' P4 wrote in his diary that he preferred minimal engagement with the apps as *"Interaction with the app requires an interruption of pace"*. P6 showed how he preferred to focus on the run and not look at his phone (V), and even unlocking a screen was noted as too annoying for P5: *"Just want to get your head down and not particularly look at the app. You just want to go straight on and get into that position."* – P6, V; *"That I need to unlock the screen to do this was a pretty major annoyance."* – P5, D.

Beyond universal usability issues uncovered through engaging with the physical nature of running in-situ,

different types of engagement were preferred by health and pleasure runners: health runners wanted a companion to motivate them and pleasure runners wanted to engage with the system on their own terms.

Information Needs Vary Based on Motivation
Self-identified motivations for running influenced information needs. Pleasure runners found post-run statistics useful and enjoyable, making them more willing to manually input information. When adding data, P5 described liking inputting post-run information for pattern recognition (CI) and P4 said he added data just in case it might be useful (CI) (*"documentary tracking"* [26]): *"Maybe you find out that you really like runs in the rain, for example, or if they have a way to pull that data together"* – P5, V; *"I never really look at them, but it's just there, just in case."* – P4, V.

Documenting the runs was not as useful to people running for health who didn't see the immediate benefit of this data collection and were *"not interested in graphs"* (P2, CI). They recalled annoyance with this extra step before finishing (Figure 4), particularly right after finishing a run that they may not have enjoyed: *"I just thought it was going to be "stop." [...] Instead, it was like, "Do this. Do that. Tell us what your run was before we save the run," and then it saves the run. That was a little bit annoying."* – P1, CI; *"I don't really want it to trigger my memory."* – P3, CI.

The preferences for the type of information differed based on their self-reported motivation, for instance health runners looking back at a rough estimate of calories burned, whereas pleasure runners dismissed this data as inaccurate: *"I think calories as well, even though I don't actually trust the calculations, but it's a*

nice rough idea." – P3, CI; *"Okay, distance and duration, average pace. Yes, I think we'll stop about here. I don't really care about everything else. I don't think it's accurate, anyways. Calories is based on your weight. I didn't even enter my weight.* – P4, V.

The findings show that information needs were very different for people who enjoyed running and those who primarily did it for health. Health runners wanted to be finished with the run as soon as possible and were only interested in information directly connected to their health goals, such as calories. Pleasure runners enjoyed having accurate post-run information available to them, whether they used it or not.

Discussion and Conclusions

There are a substantial number of user studies of PI systems that draw on different interviewing techniques (e.g. [15,26]) and mixed method approaches (e.g.[8,10]). However, few studies have taken a situated approach. Although our study method had to be adapted to the physical nature of the activity, it still revealed important user differences through observation in the context of use: motivation to exercise and the physical nature of the activity influences the use of the apps. Although there were general issues with the running app tested, self-identified health runners and pleasure runners had different expectations, engaged differently when actually running, and had different information needs post-run. Through conducting this situated study it is clear that a one-size-fits-all app is not appropriate. Building on what Consolvo et al [4] concluded with regards to customisation and adaptivity when looking at the more stationary Ubifit Garden [6], we believe that this study has implications for the design and

evaluation of future fitness technologies, including increasingly popular wearables for tracking activity.

We had initially planned to use the situated method proposed by Brown et al. [1] to understand the use of the running apps in context. However, studying strenuous physical exercise required adaptations to the methods, and we were unable to capture video of participants directly interacting with the apps. Therefore in addition to field videos captured by the experimenter, we also made use of supplementary retrospective accounts through contextual interviews and diary studies. Methods are not a binary choice between situated or not, but rather lie on a continuum between retrospective and real-time, and between 'in the wild' and laboratory. Through the triangulation of the methods used, some of which were more real-time and situated than others, we were able to capture running app use in context. While the 'in the wild' real-time methods of autoethnography and observation using field video helped uncover a wider range of usability problems, the 'in the wild' retrospective methods of contextual interviews and diary study showed issues that were most important for each participant: a narrower range of problems that were repeatedly reported, and explained reasons behind use and non-use.

Even with a small number of participants, we were able to gain insight into the stark differences in how motivations influence people's expectations, preferences, and usage of running apps in-situ. This study shows the value of a pragmatic situated approach when studying use in context by using a variety of methods, as real-time and situated as possible, to achieve triangulation and understand real world use.

References

1. Barry Brown, Moira McGregor, and Eric Laurier. 2013. iPhone in Vivo: Video Analysis of Mobile Device Use. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 1031–1040. http://doi.org/10.1145/2470654.2466132

2. Scott Carter and Jennifer Mankoff. 2005. When Participants Do the Capturing: The Role of Media in Diary Studies. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 899–908. http://doi.org/10.1145/1054972.1055098

3. Yu Chen and Pearl Pu. 2014. HealthyTogether: Exploring Social Incentives for Mobile Fitness Applications. *Proceedings of the Second International Symposium of Chinese CHI*, ACM, 25–34. http://doi.org/10.1145/2592235.2592240

4. Sunny Consolvo, Katherine Everitt, Ian Smith, and James A. Landay. 2006. Design requirements for technologies that encourage physical activity. *Proceedings of the SIGCHI conference on Human Factors in computing systems - CHI '06*, ACM Press, 457. http://doi.org/10.1145/1124772.1124840

5. Sunny Consolvo, Predrag Klasnja, David W McDonald, et al. 2008. Flowers or a Robot Army?: Encouraging Awareness & Activity with Personal, Mobile Displays. *Proceedings of the 10th International Conference on Ubiquitous Computing*, ACM, 54–63. http://doi.org/10.1145/1409635.1409644

6. Sunny Consolvo, David W McDonald, Tammy Toscos, et al. 2008. Activity Sensing in the Wild: A Field Trial of Ubifit Garden. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 1797–1806. http://doi.org/10.1145/1357054.1357335

7. Paul Dourish. 2004. What we talk about when we talk about context. *Personal Ubiquitous Comput.* 8, 1: 19–30. http://doi.org/10.1007/s00779-003-0253-8

8. Daniel A. Epstein, An Ping, James Fogarty, and Sean A Munson. 2015. A Lived Informatics Model of Personal Informatics. *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, 731–742. http://doi.org/10.1145/2750858.2804250

9. Thomas Fritz, Elaine M Huang, Gail C Murphy, and Thomas Zimmermann. 2014. Persuasive Technology in the Real World : A Study of Long-Term Use of Activity Sensing Devices for Fitness. 487–496.

10. Daniel Harrison, Paul Marshall, Nadia Bianchi-Berthouze, and Jon Bird. 2015. Activity Tracking: Barriers, Workarounds and Customisation. *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, 617–621. http://doi.org/10.1145/2750858.2805832

11. Minna Isomursu, Kari Kuutti, and Soili Väinämö. 2004. Experience Clip: Method for User Participation and Evaluation of Mobile Concepts. *Proceedings of the Eighth Conference on Participatory Design: Artful Integration: Interweaving Media, Materials and Practices - Volume 1*, ACM, 83–92. http://doi.org/10.1145/1011870.1011881

12. Shaun K. Kane, Chandrika Jayant, Jacob O Wobbrock, and Richard E Ladner. 2009. Freedom to Roam: A Study of Mobile Device Adoption and Accessibility for People with Visual and Motor Disabilities. *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility*, ACM, 115–122. http://doi.org/10.1145/1639642.1639663

13. AnaCarolinaTomé Klock and Isabela Gasparini. 2015. A Usability Evaluation of Fitness-Tracking Apps for Initial Users. In *HCI International 2015 - Posters' Extended Abstracts SE - 76*, Constantine Stephanidis (ed.). Springer International Publishing, 457–462. http://doi.org/10.1007/978-3-319-21383-5_76

14. Armağan Kuru and Jodi Forlizzi. 2015. Engaging Experience with Physical Activity Tracking Products. In *Design, User Experience, and Usability: Design Discourse SE - 46*, Aaron Marcus (ed.). Springer International Publishing, 490–501. http://doi.org/10.1007/978-3-319-20886-2_46

15. Amanda Lazar, Christian Koehler, Joshua Tanenbaum, and David H Nguyen. 2015. Why We Use and Abandon Smart Devices. *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, 635–646. http://doi.org/10.1145/2750858.2804288

16. Ian Li, Anind K Dey, and Jodi Forlizzi. 2012. Using Context to Reveal Factors That Affect Physical Activity. *ACM Trans. Comput.-Hum. Interact.* 19, 1: 7:1–7:21. http://doi.org/10.1145/2147783.2147790

17. JJ Lin, Lena Mamykina, and Silvia Lindtner. 2006. Fish'n'Steps: Encouraging physical activity with an interactive computer game. *UbiComp 2006: …*: 261–278.

18. Clara Mancini, Keerthi Thomas, Yvonne Rogers, et al. 2009. From Spaces to Places: Emerging Contexts in Mobile Privacy. *Proceedings of the 11th International Conference on Ubiquitous Computing*, ACM, 1–10. http://doi.org/10.1145/1620545.1620547

19. Matthew Mauriello, Michael Gubbels, and Jon E Froehlich. 2014. Social Fabric Fitness: The Design and Evaluation of Wearable E-textile Displays to Support Group Running. *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, ACM, 2833–2842. http://doi.org/10.1145/2556288.2557299

20. Florian Mueller and Matthew Muirhead. 2014. Understanding the Design of a Flying Jogging Companion. *Proceedings of the Adjunct Publication of the 27th Annual ACM Symposium on User Interface Software and Technology*, ACM, 81–82. http://doi.org/10.1145/2658779.2658786

21. Carman Neustaedter and Phoebe Sengers. 2012. Autobiographical design: what you can learn from designing for yourself. *interactions* 19, 6: 28–33. http://doi.org/10.1145/2377783.2377791

22. Aisling Ann O'Kane, Yvonne Rogers, and Ann E. Blandford. 2014. Gaining empathy for non-routine mobile device use through autoethnography. *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*: 987–990. http://doi.org/10.1145/2556288.2557179

23. Misha Patel and Aisling Ann O'Kane. 2015. Contextual Influences on the Use and Non-Use of Digital Technology While Exercising at the Gym. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ACM, 2923–2932. http://doi.org/10.1145/2702123.2702384

24. David Randall, Richard Harper, and Mark Rouncefield. 2007. *Fieldwork for design: theory and practice*. Springer.

25. Jens Riegelsberger and Yelena Nakhimovsky. 2008. Seeing the Bigger Picture: A Multi-method Field Trial of Google Maps for Mobile. *CHI '08 Extended Abstracts on Human Factors in Computing Systems*, ACM, 2221–2228. http://doi.org/10.1145/1358628.1358655

26. John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers Chalmers. 2014. Personal tracking as lived informatics. *Proceedings of the 32nd annual ACM conference on Human factors in*

computing systems - CHI '14: 1163–1172. http://doi.org/10.1145/2556288.2557039

27. Jakob Tholander and Stina Nylander. 2015. Snot, Sweat, Pain, Mud, and Snow: Performance and Experience in the Use of Sports Watches. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ACM, 2913–2922. http://doi.org/10.1145/2702123.2702482

28. Rayoung Yang, Eunice Shin, Mark W Newman, and Mark S Ackerman. 2015. When Fitness Trackers Don'T "Fit": End-user Difficulties in the Assessment of Personal Tracking Device Accuracy. *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, 623–634. http://doi.org/10.1145/2750858.2804269