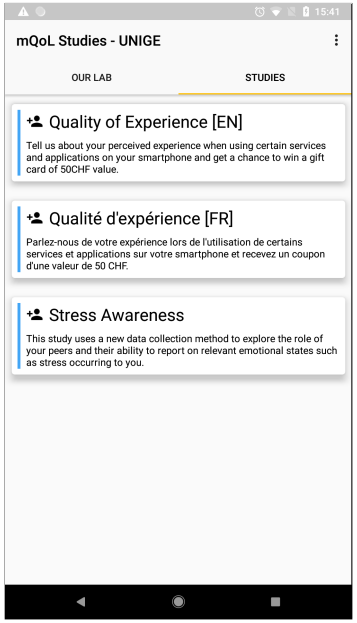

Figure 1: Study List in mQoL Lab

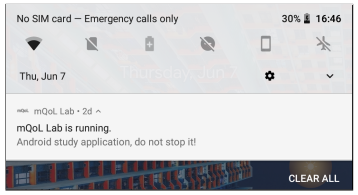

Figure 2: Running Study Notification

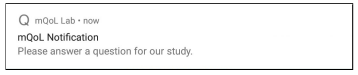

Figure 3: New EMA Notification

You're Using This App For What? A mQoL Living Lab Study

Alexandre De Masi
Quality of Life lab
University of Geneva
Geneva, Switzerland
alexandre.demasi@unige.ch

Katarzyna Wac
Quality of Life lab
University of Geneva
Geneva, Switzerland
University of Copenhagen
Copenhagen, Denmark
katarzyna.wac@unige.ch
wac@di.ku.dk

UbiComp/ISWC'18 Adjunct, October 8–12, 2018, Singapore, Singapore

ACM ISBN 978-1-4503-5966-5/18/10...$15.00.
https://doi.org/10.1145/3267305.3267544

Abstract
Smartphones are personal ubiquitous devices that provide an immense source of information via diverse applications (apps) that contribute to our decision-making process throughout the day and improve our quality of life in the long term. In the past, an app only had one or a few specific functions, while nowadays, given the same interface, an app provides multiple interactive services to their users. However, we still have a weak understanding of user expectations and experiences with these apps. Towards this end, we extended our previous smartphone logging app, to the new 'mQoL Lab' for **m**obile **Q**uality **o**f **L**ife, to strategically trigger user surveys and to achieve a better understanding of the user's actions in popular Android apps, like: Spotify, WhatsApp, Instagram, Maps, Chrome, Facebook and its Messenger. We present and discuss the results and their implications acquired during our first pilot study conducted with five users for four weeks in our Living Lab settings.

Author Keywords
Smartphone Usage, Application, Living Lab, Qualitative and Quantitative Study

ACM Classification Keywords
H.4.m [Information Systems Applications]: Miscellaneous; H.10.m [Human-centered computing]: Smartphones, Empirical studies in ubiquitous and mobile computing

Research Methods

Qualitative: (a) Entry surveys enabling us to understand the individual's socio-demographics and current experience with the smartphone/apps (b) Ecological Momentary Assessment (EMA) [13] enabling us to understand the momentary, just in context attitude, needs and behaviours of study participants using their smartphone/apps.

Quantitative: Raw and analysed datasets obtained from smartphone built-in sensors via 'mQoL Lab' app.

Hybrid: The simultaneous applications of qualitative and quantitative methods, allowing for better accuracy of datasets collected towards our end goal.

	Category
WhatsApp	Communication
Chrome	Communication
Messenger	Communication
Facebook	Social
Instagram	Social
Spotify	Music and Audio
Maps	Travel and Local

Table 1: Application Metadata from Google Play Store

Introduction

In the last 15 years, mobile phones became 'smart' and widely available, and its usage evolved into a crucial skill in supporting our day-to-day needs for information and communication "on the go".

There exists a copious amount of different applications (apps) helping us in these needs, especially via a particular app developed for a specific goal, e.g. video-chat, reading the news or playing a game. The smartphone becomes an extension of our body in our daily life, being at least 50% of the time within our arms reach [3]. It is highly probable that a smartphone (and wearables connected to it) is becoming a tool that allows us to improve our everyday activities and health, considerably contributing to our Quality of Life (QoL). We started researching the app experience in 2010 indicating the major factors influencing the popular apps [6], through challenges in human subject studies "in-the-wild" [5] and connectivity patterns of smartphone users [16], to living-lab approach to data collection in mobile studies [11], explicitly focusing on mobile users in trains or other moving vehicles. In our research, we have leveraged hybrid research methods - with different sources (the participant and his/her smartphone), different data granularity and timeliness of the data acquired from the participants. Our smartphone-based logger app, previously named mQoL-Log [15] had several limitations, e.g. it was challenging to configure, was predefined for a specific study goal and did not include context-based triggers for user surveys launched on a participant smartphone. The current, updated version, mQoL Lab, has a better study management system. It is able to run multiple human subject studies at once and engage the participants via the same interface (Figure 1).

For this study, we selected a set of the most popular interactive apps (Table 1) and asked the user to rate their experience after using these apps. Our primary goal was to test mQoL Lab on a small scale study. Our secondary goal was to understand and model the end user's Quality of Experience (QoE). Le Callet et al. [10] define QoE as *the degree of delight or annoyance of the user of an app or service"*. QoE findings are just indicated and not discussed in details in this paper due to space limitation.

Related Work

The currently available and open source research platform for human subject studies, AWARE framework[4], does not propose fine-grained tuning of context-based event triggers, specifically user survey after a specific app utilisation. Ickin et al.[6] used AWARE for their QoE study with this limitation. The survey's trigger was random during the day without a specific set of apps to rate for the user to provide their QoE rating on. Casas et al. [1] developed an app for collecting user experience in the field, where they ask their participants to execute a specific task on a set of app after which, a preprogrammed user survey appears. Most of their datasets are composed of ratings of Youtube video visualisation. We decided to discard the video apps for our work as the QoE community is heavily focused on this subject [12, 8, 9]. DeMoor et al. [2] proposed a detailed yet theoretical framework for evaluating QoE in a living lab setting in 2010. Since this framework was created, the mobile operating system landscape changed; hence it can not be entirely used today. Their solution needs profound system information not publicly available via the Android API, without installing a custom Android OS on the participant's device. Finally, Tossell et al. [14] collected smartphone-based sensing data to explore smartphone addiction, yet their solution does not make use of context-based event triggers, as we do.

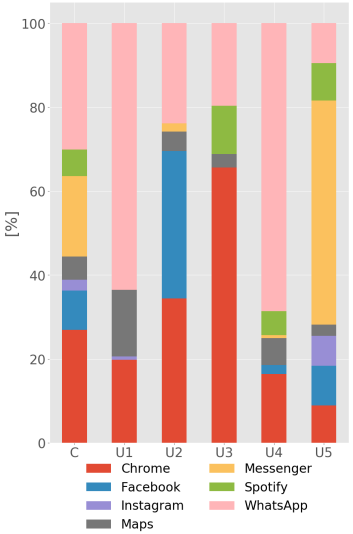

Figure 4: Distribution of Selected 7 Apps Used per User

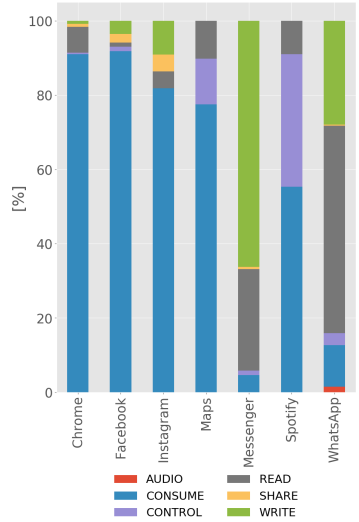

Figure 5: What action were you trying to accomplish ?

Methodology

In this study, we used a qualitative and quantitative hybrid method. The study entry survey (demographics information request and app habits) and user surveys deployed via Ecological Momentary Assessment (EMA) [13] are qualitative. EMAs allows the users to report their momentary experience with their smartphone, as the actions, they are trying to accomplish. The mQoL Lab app integrates two components, the EMA-survey manager and the 'mQoL Lab' data logger.

We have limited the scope of this study to seven popular Android apps (Spotify, WhatsApp, Instagram, Google Maps, Chrome, Facebook and Messenger (Table 1)) due to the constraint of the maximum number of EMAs to be triggered per day to minimise the participant's burden. The study has been approved by the University of Geneva's ethics commission. Overall, our study focused on expectations and the resulting QoE of mobile smartphone users. In this paper we only present results for one question that users had to reply during each EMA: *What action were you trying to accomplish?*. The user can reply to categorise his actions along seven labels: 'CONSUME content', 'SHARE or create content', 'READ text message', 'WRITE text message', 'CONTROL an app (start/stop music)', 'VIDEO call' or 'AUDIO call'. The same seven labels are always presented for any apps; it is not dynamic. It allows us to discard false replies, e.g. using Maps to make a video call.

The remaining EMA questions and possible responses are as follows:

- Did your usage of `app name` at `use start time` go as expected? Yes/No/I'm not sure.
- How was your last usage session of `app name` at `use start time`? Slider (1 to 5) with Mean Opinion Score (MOS [7]).
- Did your last `app name` at `use start time` meet your expectations? Slider (1 to 5).
- If something went wrong, please tell us more about it. Free text entry.

Once the study starts, a notification (Figure 2) is always present on the screen remembering the user that data are being collected. An EMA notification is triggered after the use of a specific app as in Figure 3, give access to a survey containing 5 questions. To further limit the user annoyance of our notifications, we set up a policy, that the mQoL Lab only trigger 12 surveys per day between the hours of 7AM to 9PM, and the minimal time between two consecutive surveys is set to 20 minutes.

Additionally to EMA, the mQoL Lab logger collects quantitative data as follows, enabling the gathering of the following information:

- Application: package name of app on the user's screen.
- Activity: user physical activity from the Google Play Services (still, tilting: between two states, in vehicle, on bicycle, on foot, running).
- Network: signal strength, basic service set identifier and service set ID (network name) of Wifi Access Point, routing tables, IP address, domain name server, ping to app server, packet and kilobytes send and received per an interface.
- Cell: signal strength, cell ID, operator name, network code and network time of the connected and neighbour cells.
- Touch: number and distribution of user touch while interacting with the screen, per a user session
- Battery: level, temperature, health and status of the battery.

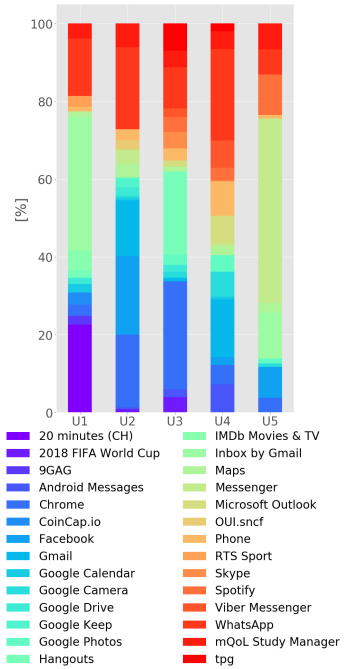

Figure 6: Distribution of Top 10 Apps per user

Recruitment

Adult participants were recruited inside the University of Geneva (Switzerland) for the duration of the study (28 days). They needed to be active users of the Android OS smartphone and users of the set of apps which we have focused on in this study; they have self-reported the top 5 apps used at the study entry time.

Results

Use of Selected Seven Apps

We have collected data for 5 users (1 female, aged 26-35, avg 32 y.o., all employed, 4 with MSc, 1 with a PhD), denoted as U1 to U5 for over 28 days. The average participation was 26 days, with an average of 7 EMAs filled per user per day. We collected 40320 minutes of cumulative data. They spend an average of 1503.6 minutes in all app session combined; that corresponds to around 5h of total app time per each participant (11 minutes/day).

From the time stamp of the app usage, we aggregate all the different mQoL Lab data sources. The cumulative distribution of the selected 7 apps usage is available in Figure 4 (**C**) as the distribution of apps per each user (**U**). One-third of our dataset is composed of WhatsApp, followed by Chrome and Messenger. The main activity during an app used is 'still'. Participants, when mobile, used their app 'on foot', followed by 'in vehicle' (tram, bus, train or car). The other activity as 'tilting' and 'on bicycle' represents a small part of our samples, presented in Figure 7.

From Figure 5, we observe the cumulative actions per selected 7 apps executed by the users. Chrome, Facebook, Instagram and Maps are more used to consume content than to share it. Messenger and WhatsApp, from the category 'Communication' (Table 1), are essentially used for reading and writing messages. Spotify user's action is shared between, consuming music and controlling the application. We conclude from it that the apps are mostly used, as they were designed from the beginning and new functionalities (like video call in WhatsApp or Messenger) are not likely to be used.

What you Say vs. What You Do: Real Use of Mobile Apps

From the overall application usage data in Figure 6, we observe that for U1, only WhatsApp is part of its top 5 apps, its top 10 integrate Maps and Chrome. U2's top 5 include WhatsApp, Facebook and Chrome, its top 10 contains Maps and Messenger. U3's top 5 incorporate Chrome and WhatsApp, its top 10 include Spotify. WhatsApp is the only app of our set in U4's top 5, and its top 10 include Chrome as well. U5's top 5 contains Messenger and Spotify, and it's top 10 include Instagram, Facebook and WhatsApp. If we compare this data with the responses from their entry survey, where they listed their top 5 apps, only U5 use all the apps that trigger an EMA. U1, U2 and U3 listed 3 of our apps set and U4 just one.

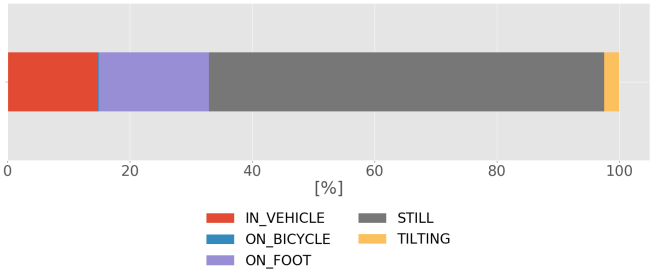

Figure 7: Distribution of Physical Activity Type (while using app)

EMA Context Survey

The mean time spent on apps is 17 ($\pm$70)[s]. We now focus on short app session (10 minutes or less, representing

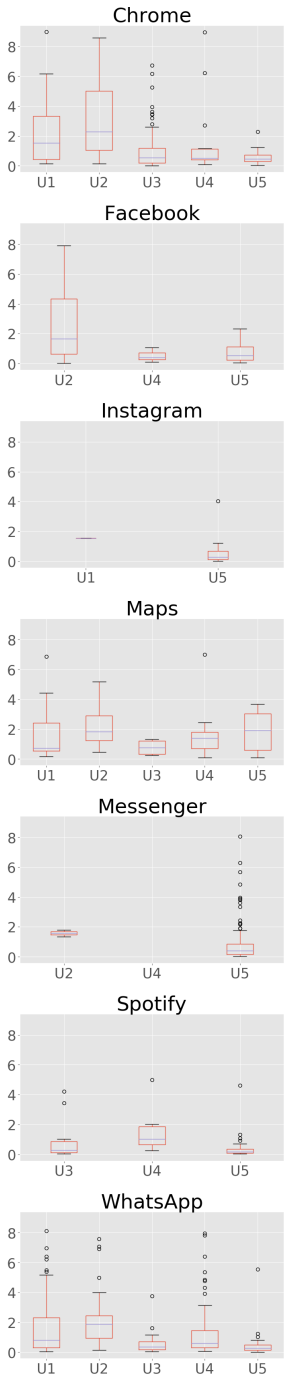

Figure 8: Time spend in each app per user [min]

89% of all sessions) and present in Figure 8 the time spent in each app by users. Chrome is the app where most time is consumed by users (22$\pm$102)[m]. Textbase conversations (WhatsApp and Messenger) are quite fast(4$\pm$30[m], 6$\pm$45[m] respectively), as choosing a song to play in Spotify (4$\pm$11[m]). The average time spent to reply to the EMA is 17.8 (± 3.9)[s]. Depending on the app, a user will spend more or less time to reply to our EMA, as presented in Table 2. We observe a high correlation (>0.8) between the expectation and experience MOS ratings.

	Mean	Std	Median
WhatsApp	4.3	119.1	10.64
Instagram	13.1	8.1	9.38
Messenger	14.2	17.3	10.1
Chrome	14.7	12.3	11.07
Facebook	16	20.05	10.04
Maps	21.3	56.2	12.11
Spotify	23	70.9	10.49

Table 2: Time spend to reply to EMA in each app [s]

Limitations
The small sample of participants is the main issue and does not allow for a total validation of the representative value of our results. We plan to open this study to a large number of people in coming weeks.

Discussion
The average number of surveys filled (7) vs maximum trigger possible (12) can be explained by two user habits. Messaging app, e.g. WhatsApp and Messenger, enables to reply to messages directly to the incoming message via built-in notification. The mQoL Lab is not able (yet) to detect this specific use case. The second habit is explained by the user app usage. Users are launching the apps in sessions, while our policy disables a notification if the previous one has been used less than 20 minutes before.

Conclusive Remarks and Future Work Areas
This study on user's actions showed the feasibility of mQoL Lab to gather interesting data relative to our goals. mQoL Lab methodological strength lies in the event-based triggering of EMA. It reduces the memory bias by firing a EMA just after the app use of interest. This mechanism may be exploited further in other studies to trigger EMA after the change of context-of interest - for example, disabled/ enabled network interface or Bluetooth, changing from indoor to outdoor, change in lighting condition. Such instrumentation will enable to closer monitor physical changes in the individual's behaviour and environment that, supported by EMA based self-reports, may help to better understand the mental, physical or state of individual. we are planning to leverage the mQoL lab in such use cases in the near future.

Acknowledgements

SNSF (157003) (2015-2019) and *AAL CoME* (2014-7-127)

REFERENCES

1. Pedro Casas, B Gardlo, and M Seufert. 2015. Taming QoE in cellular networks: From subjective lab studies to measurements in the field. *Network and Service* (2015).
2. Katrien De Moor, Istvan Ketyko, Wout Joseph, Tom Deryckere, Lieven De Marez, Luc Martens, and Gino Verleye. 2010. Proposed framework for evaluating quality of experience in a mobile, testbed-oriented living lab setting. *Mobile Networks and Applications* 15, 3 (2010), 378–391.
3. Anind K Dey, Katarzyna Wac, Denzil Ferreira, Kevin Tassini, Jin H Hong, and Julian Ramos. 2011. Getting

closer: an empirical investigation of the proximity of user to their smart phones. *The 13th international conference on Ubiquitous computing* (2011).

4. Denzil Ferreira, Vassilis Kostakos, and Anind K. Dey. 2015. AWARE: Mobile Context Instrumentation Framework. *Frontiers in ICT* 2, April (2015), 1–9.

5. Mattia Gustarini, Selim Ickin, and Katarzyna Wac. 2013. Evaluation of challenges in human subject studies "in-the-wild" using subjects' personal smartphones. *ACM conference on Pervasive and ubiquitous computing adjunct publication - UbiComp '13 Adjunct* (2013), 1447–1456.

6. Selim Ickin, Katarzyna Wac, Markus Fiedler, Lucjan Janowski, Hong Jin-Hyuk, Anind K Dey, Jin-Hyuk Hong, and Anind K Dey. 2012. Factors influencing quality of experience of commonly used mobile applications. *Communications Magazine, IEEE* 50, April (apr 2012), 48–56.

7. ITU-T Recommendation P.800.1. 2016. Mean Opinion Score Terminology. (2016).

8. Parikshit Juluri, Venkatesh Tamarapalli, and Deep Medhi. 2015. Measurement of Quality of Experience of Video-on-Demand Services: A Survey. *IEEE Communications Surveys Tutorials* 18, c (jan 2015), 401–418. DOI: http://dx.doi.org/10.1109/COMST.2015.2401424

9. István Ketykó, Katrien De Moor, Toon De Pessemier, Adrián Juan Verdejo, Kris Vanhecke, Wout Joseph, Luc Martens, and Lieven De Marez. 2010. QoE measurement of mobile YouTube video streaming. *Proceedings of the 3rd workshop on Mobile video delivery - MoViD '10* (2010), 27. DOI: http://dx.doi.org/10.1145/1878022.1878030

10. Patrick Le Callet, Sebastian Möller, and Perkis Andrew. 2012. Qualinet White Paper on Definitions of Quality of Experience. *European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003)* March (2012).

11. Alexandre De Masi, Matteo Ciman, Mattia Gustarini, and Katarzyna Wac. 2016. mQoL smart lab: quality of life living lab for interdisciplinary experiments. *UbiComp Adjunct* (2016), 635–640.

12. Hyunwoo Nam, Kyung Hwa Kim, and Henning Schulzrinne. 2016. QoE matters more than QoS: Why people stop watching cat videos. In *IEEE INFOCOM*, Vol. 2016-July.

13. Arthur A Stone and Saul Shiffman. 1994. Ecological momentary assessment (EMA) in behavorial medicine. *Annals of Behavioral Medicine* (1994).

14. Chad Tossell, Philip Kortum, Clayton Shepard, Ahmad Rahmati, and Lin Zhong. 2015. Exploring smartphone addiction: insights from long-term telemetric behavioral measures. *International Journal of Interactive Mobile Technologies (iJIM)* 9, 2 (2015), 37–43.

15. Katarzyna Wac, Gerardo Pinar, Mattia Gustarini, and Jerome Marchanoff. 2015a. More mobile & not so well-connected yet: Users' mobility inference model and 6 month field study. In *2015 7th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*. IEEE, 91–99.

16. Katarzyna Wac, Gerardo Pinar, Mattia Gustarini, and Jerome Marchanoff. 2015b. Smartphone users mobile networks quality provision and VoLTE intend: Six-months field study. In *IEEE 16th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*. IEEE, 1–9.