TimeUse+
Testing a novel survey for understanding travel, time use, and expenditure behavior

Working Paper

Author(s):
Winkler, Caroline; Meister, Adrian; Schmid, Basil; Axhausen, Kay W.

Publication date:
2022-08

Permanent link:
https://doi.org/10.3929/ethz-b-000561733

Rights / license:
In Copyright - Non-Commercial Use Permitted

Originally published in:
Arbeitsberichte Verkehrs- und Raumplanung 1767

Funding acknowledgement:
189001 - Consumption and travel after the smartphone revolution (SNF)
TimeUse+: Testing a Novel Survey for Understanding Travel, Time Use, and Expenditure Behavior

Caroline Winkler
Institute for Transport Planning and Systems (IVT) ETH Zürich, 8093 Zürich, Switzerland
Email: caroline.winkler@ivt.baug.ethz.ch
ORCiD: 0000-0002-0733-8990

Adrian Meister
IVT ETH Zürich, 8093 Zürich, Switzerland
Email: adrian.meister@ivt.baug.ethz.ch
ORCiD: 0000-0002-3350-9044

Basil Schmid
IVT ETH Zürich, 8093 Zürich, Switzerland
Email: basil.schmid@ivt.baug.ethz.ch
ORCiD: 0000-0002-3310-9083

Kay W. Axhausen
IVT ETH Zürich, 8093 Zürich, Switzerland
Email: axhausen@ivt.baug.ethz.ch
ORCiD: 0000-0003-3331-3138

Word Count: 6728 words + 3 table(s) × 250 = 7478 words
Submission Date: August 1, 2022
ABSTRACT

Travel diaries are commonly used to better understand how and why residents of an area choose to move around their environment. Recent studies have increased the complexity of their travel diaries, be it through accounting for further relevant characteristics within the diary itself or with questionnaires, all with the goal of accurately contextualizing these travel decisions. TimeUse+ is a novel tool that collects GPS tracks passively and has users actively enrich tracked events with activity and expenditure information all within the same diary app. We document all components of our flexible platform in their entirety before evaluating the data collection tool in two applications with over 200 individuals in Switzerland. We test several configurations that alter study duration, detail of activities performed, and the monetary incentive offered. Results suggest that it is reasonable to apply our framework for an extended period of four weeks, yielding rich longitudinal accounts of individuals’ travel, time use, and expenditure patterns. We discuss adjustments made along the way to improve the app and accompanying materials and present initial results about how Swiss residents use their time and money. We ultimately describe limitations of the framework and outline next steps for the project.

Keywords: smartphone survey, GPS tracking, travel behavior, time use, household expenditures
INTRODUCTION
Travel diaries are widely implemented to understand travel behavior as they allow for trips to be considered within the broader context in which they take place. This of course includes the activity that is undertaken at the destination, but also characteristics of the individual performing the trip. More realistically, the activity chain of an entire day, household structure and division of responsibilities, monetary resources available, and general lifestyle and values will all affect the choice to make a trip. The goal of collecting a comprehensive data set for travel behavior research has been shared by other recent studies, like Calastri and colleagues (1) who used a life-course calendar and surveys to capture social relationships along with travel and activity information. Not only should the entire context be taken into consideration, but at least two weeks of an individual’s travel behavior need to be observed to properly capture their variability, habits and true activity space sizes (2, 3, 4). With so many variables of interest, it is imperative that any tool that aims to collect such data does so without causing much burden for the respondents.

We took on the challenge of integrating the three crucial components of time use, travel, and expenditure data collection into a single mobile application, the first striving to do so over the course of an entire month. Such novel methods require extensive testing before they can be applied at a national scale. This paper shares the experience gained through performing the TU+ study with over 200 individuals in German-speaking Switzerland. This evaluation focuses on four important aspects: 1) are individuals willing to participate in the study over an extended period of time?, 2) under what app configuration is this most likely?, 3) does the study capture the three realms of interest as expected?, 4) how do Swiss individuals actually spend their time and resources?

First, we review existing literature on survey methodology related to travel diaries and broader literature on participation in smartphone data collection. The survey design and components are then described in detail before presenting the procedure and experience recruiting 220 participants for a pilot and pretest. A discussion of the results reveals how realistic the use of the TU+ app is before uncovering the behavioral patterns of the Swiss population. The final section provides a summary of the findings and the next applications of the TU+ study.

LITERATURE REVIEW
Often performed by governmental agencies, national surveys that aim to understand how the population uses their time, money, and how they choose to travel have been around for decades. In their simplest form, they are censuses, tools that collect data that we need to estimate and predict mobility trends, plan infrastructure developments and services, and deal with societal inequity. In Switzerland, the Federal Statistical Office manages two such studies: the Household Budget Survey performed continuously with around 250 new households each month who document every purchase they make (5), and the Swiss Mobility and Transport Microcensus (MTMC) (6), a telephone interview that covers a representative sample of 56,000 individuals’ travel behavior for a single day, including trip purpose. National time use surveys are conducted in most European countries using guidelines put forward by the statistical office of the EU (Eurostat) to harmonize responses across countries. Here, participants keep a time use diary for one weekday and one weekend day with everything they have been doing in 10-minute intervals, including secondary activities, locations, and social partners (7). Switzerland does not administer a national time use survey.

We are at an age where the only reasonable medium for such data collection is a smartphone, since it rarely leaves people’s sides (especially when outside of the home), and 92% of
adults in Switzerland own a smartphone (97% of whom use it every day). Though the incorporation of passive logging with a smartphone into travel diary studies has been around for over 10 years, the true potential of using smartphones as a data collection tool is arguably only beginning. Issues of battery consumption and GPS sensors/coverage are perhaps the main concerns mentioned when doing so (e.g.,). Today, though quality of battery life and GPS sensors may not be drastically better, smartphone apps can be programmed in a way that limits battery usage (e.g. by only running the app when the phone is in motion).

Accuracy concerns as they relate to automatic transport mode detection are the next prominent issue in recent smartphone studies. Harding and colleagues recently compared 17 different tracking apps on 21 phones while manually recording ground-truth data and found that detected modes across the devices never to be more than 75% accurate. A way of dealing with this is therefore adding an active (and subjective) diary component to correct automatically detected transport modes. That being said, the MOBIS study, which uses MotionTag mode detection, had accuracy rates around 92% for both iOS and Android devices. Lynch and colleagues highlight the increases in data quality that can be achieved using passive GPS diaries; they found more detailed trip time, distance, duration, path/location information compared to an online/call center diary. What is more challenging to understand from these objective tracks is why a trip was made. Trip purpose can be imputed from GPS tracks using machine learning techniques, as is often done (e.g.,), but a subjective account is necessary to confirm these data and to generate training data.

Moreover, the travel behavior realm is interested in an accurate reporting of trip purpose because it is relevant for traditional forecasting models of mode and route choice. Reliable travel diary data also serve transport planning and policy by facilitating the estimation of time value metrics like that of value of travel time (VTT), which finds itself in virtually all related cost-benefit analyses. A better understanding of what lies behind VTT requires not only instances of travel behavior, however, but also of how much time a person spends doing other (un)paid work and leisure activities, and information on how that person spends their monetary budget. It is for that reason that many travel diary studies of today collect time use information, though unfortunately usually in the simple form of trip purpose. Naturally, multiple activities can be undertaken at a location, which is often the case, and how much time is spent on each activity provides valuable information.

Almost all travel diaries to date have left out expenditure-related aspects necessary for value of leisure estimations in their study design. To date, only two paper-and-pencil diary studies, Mobility-Activity-Expenditures-Diary (MAED) and Post Car World (PCW) have managed to do so, as well as one recent smartphone diary. All three studies ran over only 7 days and the smartphone study by Alho and colleagues only considered daily expenses and used an app exclusively available for Android devices. Asking for these active inputs of activities and expenditures naturally leads to increases in burden for respondents, and in turn, potential for increased fatigue over time.

TIMEUSE+ PLATFORM

The TU+ platform consists of two components which are illustrated in Figure. The TU+ app is the core element of the platform and the tool used by participants to record their joint travel, time use, and expenditure diaries. The app uses the Software-Development-Kit (SDK) from MotionTag that sends sensor data to the external MotionTag API, which in turn performs the trip and mode
detection before communicating the results to the TimeUse+ backend. The diary format makes use of two different event types, tracks and stays, to fill a respective 24 hour timeline. Each event can be enriched with activities and further contextual information, including social partners present and amount of money spent. Our integrated diary format is constructed within the TU+ API by merging the travel diary from MotionTag with time use and expenditure related diary components that are self-reported by the user. This merged data-stream is stored in a database, and can be fetched from the app for visualization. Furthermore, the app accesses a configuration file through the TU+ API that specifies various generic aspects of the survey format. This configuration file can be edited by the researcher through the WebAdmin (described below).

![FIGURE 1: TimeUse+ platform architecture](image)

**TimeUse+ Webadmin**
Our platform allows for event types, activity types and activity attributes of interest for a given study to be set by the researchers in the WebAdmin (online interface). These are dynamically retrieved through the TU+ API using the configuration file. One can for instance incorporate certain transport modes to represent a regional transport mode-mix or determine the list of activities and expenditures according to official guidelines. The list of activities presented for each event type is configured separately, which facilitates the users’ experience, as they are not presented unlikely combinations (e.g. sleeping at the workplace). This also permits researchers to investigate combinations of interest to specific research questions like online shopping during travel.

**TimeUse+ App**
The left screenshot in Figure 2 represents the home screen that appears when a (logged-in) user opens the app. The control bar on top of the screen can be used to open the calendar view to change the selected day, and includes buttons to open statistics and settings screens. The calendar overview visually indicates which days still require validation (i.e. reporting of activities and expenditures) by the user. The settings screen allows the user to access general settings, the FAQ section, and to control battery saving mode. The statistics screen summarizes the travel and time use behavior for a given week, showing the most performed trips/activities as well as generated CO2 emissions for the former. The home screen displays the timeline of a tracked day with a corresponding map.
The timeline shows all recorded events with their start and end time, and visually differentiates between stays and trips and whether an event has been validated or not. Events up to five minutes long are automatically validated, that is, a respondent is directly shown a blue check mark and a greyed-out "Edit" button instead of the "Add Details" option. Placeholder events are inserted when there is a gap in the timeline due to malfunctioning tracking to ensure the continuity of the timeline and not annoy the user with missing time slots. The timeline therefore spans a full 24 hours starting at midnight each day. However, the first and last listed event typically stretch into the previous or next respective day for a better overview and to simplify the validation task for users who can then see the diary information they recorded on both days.


When tapping a specific event in the timeline, the map’s focus changes accordingly and the user can open a separate screen with the details of that event, shown in the middle of Figure 2. The Event Details screen shows a map with the detected location(s) of an event. Users can correct wrongly detected modes or locations by clicking on the location/mode icon which opens a list from which the user can choose from. Three dots to the right of the location/mode icon allow for a user to merge the entire event to the previous event or to indicate that the event did not happen at all. The Event Details screen also includes a list of activities that can be selected by the user and ends with a button that allows users to delete a single event, in line with data and privacy regulations.

Finally, the screenshot at the right of Figure 2 shows the Activity Details screen that opens as soon as an activity is chosen on the Event Details screen and is used to specify attributes like duration, social partners and expenditures. Only the duration of each activity is to be reported.
by the user. The order (and therefore timing) of activities are not requested as part of this study. Duration options are given in 10 minute intervals, as is the case in typical household time use diaries. The maximum duration is naturally the maximum event duration. The range and interval for the expenditures are similarly requested in 10 CHF intervals up to 500 CHF, which is set by the researchers in the WebAdmin. The options we defined for social partners include being alone, with household members, or with friends, though the question itself and its options are fully flexible and configurable through the WebAdmin as well. Equivalently, the expenditure attributes were specifically set for our use case, and therefore ask for whether expenses were personal or for the household, daily or long-term needs, and "required" purchases or not, all of which are explained to respondents during study introduction. Again, because of the flexibility of all of the platform’s components, we sparingly ask participants for expenditures and expenditure attributes only at locations and for activities where money is likely to be spent, such as online shopping.

STUDY DESIGN

Pilot study

All components of the TU+ project were tested for the first time in May 2021 with 15 individuals. Nine hundred invitations were sent to individuals residing in German-speaking Switzerland, meaning that the net response rate was slightly more than 1.5%. The initial questionnaire was, however, completed by 35 individuals, four of whom indicated that they were not interested in participating in the two succeeding portions of the study. When asked why they did not wish to participate, two individuals stated the time effort was too high, while the other two reported not wanting to be tracked.

Participants were invited to participate with a letter by mail that included a link to an online questionnaire that described the three components of this pilot: I. the initial questionnaire, II. seven days of tracking and validating using the TU+ app, and III. participating in a video interview. The interview was semi-structured and arranged to collect information on participants’ perception of the following dimensions: onboarding, usability, and statistics screen in the TU+ app, the provided manuals and FAQs, as well as the communication from the research team’s side and motivation for participation. This pilot served as a proof of concept. These were the main takeaways:

- Onboarding: No issues were posed by the initial questionnaire, downloading and registering within the app, and for the most part completing the one-week tracking and validating period. Some respondents mentioned it not being immediately clear how validation (i.e., the active time use and expenditure component) works. This led to the Validation flyer. It also took users longer than expected to complete, around 10 minutes daily instead of the five we assumed.
- Usability: The app was regarded as generally highly usable. How activity duration was presented to and documented by the user was not intuitive, something we changed for the pretest.
- Statistics screen: Most participants reported not using it or even knowing it existed. Those particularly interested in tracking their mobility patterns found it appealing and useful.
- Supplementary materials: The FAQs and manuals were all perceived as adequate, but were rarely used. The Quick start flyer was deemed helpful.
- Communication (in-app daily notification and informational emails): Experienced as sufficient, appropriate, and not intrusive when acknowledged, but ignored by many.
- Motivation: No one participated purely for the money. Top reasons were: interest in mobility
patterns and tracking studies, curiosity, wanting to contribute to research because of the societal impact.

- General: Initial participation rate can (unsurprisingly) be expected to be low. The overall complexity of the study is not too difficult or burdensome for respondents (at least for one week) and none of the 15 users had a problem disclosing/recording their daily expenditure information. Two-thirds of participants even said they would participate again if the study ran over two or four weeks. Participants did request explanatory videos and more background information about the study itself, how their data will be used, and why it is important that they participate.

**Pretest**

**Configuration**

The procedure for our pretest was again made up of three components: an initial questionnaire, the tracking period using the TU+ app, and a final questionnaires after successfully completing the tracking period. The main goal of the pretest was to test different configurations of the study’s flexible components: participation duration, granularity of the activities participants use to describe their days, and incentive levels. All this to evaluate how long we could realistically keep participants engaged, which list of activities users would find easier to fill out, and whether it would make sense to double the incentive (i.e., would the response rate grow accordingly)? We consequently defined five different groups:

1. 4 weeks tracking & validating / detailed activity list / 50 CHF incentive
2. 4 weeks tracking & validating / simple activity list / 50 CHF incentive
3. 2 weeks tracking & validating / detailed activity list / 50 CHF incentive
4. 2 weeks tracking & validating / simple activity list / 50 CHF incentive
5. 4 weeks tracking & validating / detailed activity list / 100 CHF incentive

Participants were recruited using our invitation letter by mail. A random sample of addresses in the German-speaking part of Switzerland (18+) were bought from a marketing service provider. Of the 7,500 addresses acquired, 1,500 were used for each test group. The activity list complexity was tested on the notion that a shorter list of activities means participants have less to scroll through, and thereby less burden. The simple activity list is an aggregated version of the detailed one. For example, while the simple activity list only had "leisure" as an activity at location "other", the detailed list had entertainment, exercising, gastronomy, and socializing.

**Communication Protocol**

Figure 3 shows a schematic overview of the protocol defining how we communicate with the participants, i.e. at what time and using which materials. The invitation letter again included a QR code and web address for the initial questionnaire, and the only difference between groups’ letters were the characteristics of participation duration and incentive, according to the groups defined above. The initial questionnaire collects personal and household-level socioeconomic information, including mobility tool ownership. The questionnaire is rather standard and similar to that used in PCW and MAED (17, 16). Participants were not made aware of the list of activities they were presented in the app. The letter has the logo of the study, but also of the university conducting the study, which is highly esteemed in the country. Again, this initial sign of trust is very important for recipients, and likely increases their willingness to participate (19). The QR code was very popular, about 82.5% of pretest participants used it instead of the web address.
If respondents were eligible and confirmed wanting to participate, they provided their email address at the end of the initial questionnaire and were then forwarded to a flyer on our website (see Figure 4 left). The flyer instructs participants how to download and install the app, but most importantly, how to configure the respective phone settings so that the app starts tracking properly. Apart from the flyer, the website serves as an additional platform where participants can find more information about the study and pull up important supplementary materials like the FAQs and the app’s manual/user guide. If participants have further questions, they are urged to contact the research team via email or phone.

Once respondents finished the initial questionnaire, they were addressed using automated email communication. After successfully installing the TU+ app, participants received a confirmation email which included a flyer, similar to the one shown in Figure 4 left, explaining how to validate the recorded data, specifically how to report different activities. Participants who did not download the app within two days were sent a reminder email. A further reminder was sent out if participants downloaded the app, but no recordings were detected. The tracking period started...
with the first recorded event and then lasted for 14 or 28 days depending on which group they were assigned to. During the tracking period, participants were sent validation reminders by email every fourth day if the total validated time was below 70%. Additionally, the app triggered a daily push-notification to kindly remind participants to validate their data.

Apart from the incentive, participants were sent personalized reports in order to keep those motivated and reduce attrition. A sample report is shown in Figure 4 on the left. It was sent out once a week and includes a travel and a time use section, both showing statistics of the five most performed trips or activities, respectively. The report further compares the shares of each type of trip/activity behaves to those of the previous week, as well as to the rest of the study participants.

The statistics about the travel and time use behavior were additionally accessible directly in the app through a dedicated statistics screen.

If participants reached the required amount of tracking with a validation of at least 70%, they were directly sent a link to the final questionnaire, i.e. the very last part of the TU+ study. The final questionnaire collects long-term expenditure information, that is, expenses that essentially do not occur on a day to day basis. They include rent, insurance, vacation expenses, and monthly services and subscriptions. We additionally ask respondents about their experience with the study and the TU+ app before asking for their banking information so that they may receive their monetary incentive. Both the initial and final questionnaires are Qualtrics web-questionnaires.
that require about 10 to 15 minutes each to complete and are mobile-friendly. Participants who did not fill out the final questionnaire, were sent a reminder after 2 days. Participants that had validation lower than 20% were excluded from the last part of the study. They were sent an email reminding them that validating the data is a requirement to successfully complete the study. This was done to prevent participants from validating missing data last minute in order to be eligible for the compensation. Participants that had validation over 20% but under 70% were sent an email reminding them of the just stated requirements. They were given a grace period of 2 additional days to "complete" the validation, without stating our internal threshold of 70%. After that grace period, participants were either sent the link to the final questionnaire (with respective reminder after 2 days), or excluded analogue to those having less than 20%.

RESULTS AND DISCUSSION

The results of the pretest can be summarized by returning to the initial questions of interest.

Are individuals willing to participate in the study over an extended period of time?

A first indication of whether this is the case can easily be assessed by looking at the rates of participation and attrition at different stages of the study. As can be seen in Table 1, each group had between 6 and 10% of invitees complete the initial questionnaire, and the net response rate lies between 2.2 and 3.7%. This suggests that participants were indeed willing to participate, and though the net response rate seems low, it is comparable to recent GPS tracking studies (e.g. (12, 1)). This is exciting considering that earlier studies usually run for only a single week, and 127 of our participants successfully completed the TU+ study as it ran over 4 weeks. TU+ is also more complex because of the activity and expenditure attributes it requests from participants. Our strategy of keeping participants engaged by readily helping them along the way and providing engaging personalized weekly reports seems to at least in part serve its intended purpose of minimizing fatigue.

<table>
<thead>
<tr>
<th>Group</th>
<th>Initial Q</th>
<th>Tracked</th>
<th>Completed</th>
<th>Final Q</th>
<th>% Final Q/Initial Q</th>
<th>% Invited/Completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>126</td>
<td>77</td>
<td>50</td>
<td>39</td>
<td>30.95</td>
<td>2.60</td>
</tr>
<tr>
<td>2</td>
<td>92</td>
<td>59</td>
<td>49</td>
<td>33</td>
<td>35.87</td>
<td>2.20</td>
</tr>
<tr>
<td>3</td>
<td>124</td>
<td>84</td>
<td>75</td>
<td>40</td>
<td>32.26</td>
<td>2.67</td>
</tr>
<tr>
<td>4</td>
<td>122</td>
<td>79</td>
<td>57</td>
<td>38</td>
<td>31.15</td>
<td>2.53</td>
</tr>
<tr>
<td>5</td>
<td>157</td>
<td>103</td>
<td>76</td>
<td>55</td>
<td>35.03</td>
<td>3.67</td>
</tr>
<tr>
<td>sum</td>
<td>621</td>
<td>402</td>
<td>307</td>
<td>205</td>
<td>mean 33.05</td>
<td>2.73</td>
</tr>
</tbody>
</table>

Note: M1-M4 refer to the milestones outlined in Figure 3. Q stands for questionnaire and Final Q reflects the final sample. Tracked individuals are those who began tracking using the TU+ app and Completed means that they reached the end of their required tracking period, but not necessarily that they validated the required 70% of events.

Which configuration of the study design works best?
Table I does not clearly indicate that one configuration may be better than the other. In contrast, email communication and troubleshooting related to activity classification was experienced by the research team as much more frequent for participants of groups 2 and 4 (simple list). Though our original hunch was that additional scrolling would cause more burden, it (quite reasonably) seems
that it is much easier for respondents to fill out their diary when they see exactly the activity they were just performing and do not have to sit and think about how to categorize an activity when the list has fewer, but broader options.

As far as having validated data, all groups validated at least 95% of their days. This is unsurprisingly the case because it was a study requirement and participants were not made aware that our threshold was set lower at 70%\footnote{This was purely done so that the research team would not prematurely and inadvertently kick individuals out of the study who did not directly validate their events and would therefore have a lower validation percentage for a few days during the study period. Weekly reminders were sent to participants to either gently scold them for not validating enough or thank them and ask them to keep going.} A second way we evaluated data quality was by comparing the number of activities tracked for each day, which was similar across all groups at around 17 activities per day.

Certainly, having such rich data from 205 individuals is invaluable, however, given that 70% of people who first started tracking and validating did not make it to the end, it is imperative to soon ensure that there are no systematic biases in the final sample. If we return to consider the response rates, an important finding is that there was no large difference between those who tracked two weeks and four. It seems that burden does not grow linearly with tracking duration; if someone was already willing to keep a diary for two weeks, two more was not too much to ask. Though the net response rate was indeed highest for Group 5 (100 CHF incentive)—and this by around 40%—that group was also uniquely offered a monetary incentive twice as high as the four other groups. It cost the research team almost 3 times more in incentives for only 16 more individuals.

**Does the study capture the three realms of interest as expected and how do the Swiss spend their time and resources?**

Data were aggregated for analyses as quality and response rates did not differ greatly between groups. The 205 participants tracked and validated a total of 4,480 days, 21.85 days each, on average. Participation started on March 3rd, 2022 and concluded on April 27th. Fifty percent of tracked days occurred before March 20th. An average daily total time of 9.05 hours were spent at home, 11.19 hours at an "other" location, 3.06 hours at work, and 2.19 traveling over all days and for all participants. The travel duration is exceptionally high because some participants were tracking over Easter weekend that fell on April 9th. If only the first month of the study period is considered, the average amount of time spent on travel each day is 1.63 hours.

**Travel**

As mentioned, travel behavior was passively recorded by the app. 32,335 stages were recorded over the course of two or four weeks from 205 participants. Each stage, with its automatically detected mode, was presented to participants who then corrected the mode, if necessary, and added activities if the stage was at least five minutes long. Table further shows the share of trips made with each of the reported modes along with their average distances. The full list of modes was collapsed into the following categories for analyses:

- Walk
- Bicycle: private (e-)bicycle
- Local PT: bus, tram, subway
- Train: (regional) train
• Car/motorbike: private car as driver or passenger, motorbike
• Other: carsharing, taxi/Uber, kick scooter, (e-)bicycle sharing, airplane, boat, coach, cable
car, ski, other

TABLE 2: Travel behavior overall by mode

<table>
<thead>
<tr>
<th>Mode</th>
<th>% Mode share</th>
<th>Mean distance (km)</th>
<th>SD distance (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walk</td>
<td>50.44</td>
<td>0.96</td>
<td>3.13</td>
</tr>
<tr>
<td>Car/motorbike</td>
<td>29.72</td>
<td>13.57</td>
<td>24.30</td>
</tr>
<tr>
<td>Local PT</td>
<td>7.20</td>
<td>3.20</td>
<td>3.14</td>
</tr>
<tr>
<td>Bicycle</td>
<td>6.20</td>
<td>3.38</td>
<td>4.29</td>
</tr>
<tr>
<td>Train</td>
<td>4.55</td>
<td>29.48</td>
<td>44.31</td>
</tr>
<tr>
<td>Other</td>
<td>1.89</td>
<td>45.46</td>
<td>444.59</td>
</tr>
</tbody>
</table>

Note: PT = public transport, SD = standard deviation

Time use
Time use data arise from the activities manually reported by participants at fixed locations and while traveling. Time use at the fixed locations of home, other, and work, as described above, were detected 3,789, 17,707, and 3,411 times, respectively. This means that over 20% of tracked days did not have a stay event at home. This is unlikely, and presumably due to the app’s default location being “other”. The app does learn home and work locations over time, but especially at the beginning of study participation, it may not yet be detected as a participant’s home location and/or it could be unclear to participants that they can correct the location type.

A descriptive analysis of participation in different activities and average amount of time doing said activity (given participation) is presented in Table 3. Participation in activities for these stay events is always with respect to the location in which it took place since the list of activities was location-dependent. For the sake of presentation, activities were collapsed into the broader categories of the simple activity list. For home location, only stays that were at least five hours long are included in the table to show how people spend their time during the longer stints at home instead of the two hour stops that happen throughout the day. For "other" location, only events at least five minutes long (that require activity info) are included.

The aforementioned unfortunate inaccuracy in home location is also reflected in the activity participation rates shown in Table 3 for location other. We found relatively high rates of participation in activities like eating/cooking, self-care, and sleeping, while the most likely activities are rather leisure and shopping for "other" stays. Fortunately, this issue can potentially be rectified by examining participants’ GPS tracks, identifying their home locations spatially, and relabeling them accordingly.

Most time spent on activities at home is in line with expectations. Virtually all events at home had a self-care and sleep activity, for instance, and participants reported sleeping for an average of around 7.5 hours per night. One would have expected eating/cooking to also take place during close to 100% of these events because one usually eats dinner and/or breakfast while at home for this time period, but it could also be that one truly only gets ready for bed and goes to sleep during a busy week. Leisure was also often undertaken while at home, as was digital entertainment and doing chores. Interestingly, a quarter of the events at home had some kind of...
working or studying activity performed for an average of almost 3 hours. Working from home was definitely present and common for at least some participants in our pretest. Somewhat unfortunate is how often the activity "other" was performed at each location (and for how long). It is impossible to tell whether participants were unwilling to take the time to correctly validate these events or if they genuinely did not know how to categorize certain activities.

**TABLE 3: Time use behavior**

<table>
<thead>
<tr>
<th>Location</th>
<th>Activity</th>
<th>% participation</th>
<th>Mean duration (min.)</th>
<th>SD duration (min.)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Home</strong></td>
<td>Chores</td>
<td>60.49</td>
<td>92.32</td>
<td>82.93</td>
</tr>
<tr>
<td></td>
<td>Digital entertainment</td>
<td>60.32</td>
<td>129.53</td>
<td>85.08</td>
</tr>
<tr>
<td></td>
<td>Eating/cooking</td>
<td>82.50</td>
<td>85.91</td>
<td>49.30</td>
</tr>
<tr>
<td></td>
<td>Leisure</td>
<td>71.24</td>
<td>100.69</td>
<td>78.20</td>
</tr>
<tr>
<td></td>
<td>Online shopping</td>
<td>6.12</td>
<td>41.15</td>
<td>22.88</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>22.62</td>
<td>135.36</td>
<td>170.76</td>
</tr>
<tr>
<td></td>
<td>Self-care</td>
<td>96.09</td>
<td>62.38</td>
<td>64.64</td>
</tr>
<tr>
<td></td>
<td>Sleeping</td>
<td>94.43</td>
<td>463.57</td>
<td>99.05</td>
</tr>
<tr>
<td></td>
<td>Working/studying</td>
<td>25.47</td>
<td>176.15</td>
<td>153.53</td>
</tr>
</tbody>
</table>

| **Other** | Eating/cooking     | 45.66           | 97.39               | 68.03              |
|           | Errands            | 18.39           | 172.06              | 151.14             |
|           | Leisure            | 70.25           | 210.65              | 195.59             |
|           | Other              | 19.19           | 292.80              | 279.91             |
|           | Self-care          | 42.70           | 124.32              | 193.64             |
|           | Shopping           | 6.58            | 78.08               | 94.89              |
|           | Sleeping           | 68.27           | 521.73              | 161.72             |
|           | Waiting            | 4.62            | 107.82              | 198.27             |
|           | Working/studying   | 32.68           | 314.90              | 187.49             |

| **Work**  | Other              | 32.93           | 201.75              | 302.79             |
|           | Working            | 97.54           | 465.13              | 148.05             |

**Expenditures**

Expenditures that are large and infrequent—what we call long-term expenditures—were collected during both the initial and final questionnaires, while daily spending was recorded directly in the app.

Data collected through questionnaires reveal participants’ household income, how finances in the household are distributed, and typical vacation spending patterns. In order to infer whether daily purchases that accumulate throughout a day were adequately collected, we investigated daily expenses like shopping, leisure, and errands at location “other” reported subjectively. It is also the case here that the incorrectly labeled/validated “other” stay events skew results to an extent, which is evident in Table [3] as only 6.58% of other events include a shopping activity. On average, for any given week, those who participated in said activities spent 8.23 CHF on eating/cooking, 10.38 CHF on errands, 23.30 CHF on leisure, and 62.23 CHF on shopping. These rates are not in line with expectations, and are substantially lower if all participants who tracked each week are included. A more detailed examination of these subjectively recorded expenses is required, but there is indeed evidence that participants did not diligently record their expenditures (reasons for this behavior are comprehensively discussed by Schmid [14]).
SUMMARY

This paper describes and evaluates test phases of the TimeUse+ study. The elements of the entire platform are explicitly documented before presenting results from preliminary tests of the survey tool completed with over 200 individuals for whom two or four weeks of diary information was collected. Results show that it is possible and reasonable to run the TU+ study at a large scale with a high complexity of information requested from participants.

A few limitations of the study can be identified. Mobile app development is firstly a challenge that should not be underestimated, and ever-changing requirements set by Google and Apple only increase the required efforts. MotionTag, whose SDK the app is built on, is always updated accordingly, but whether the TU+ app should be updated each time is unfortunately rather evaluated on the basis of urgency. It costs both time and money to have changes done to the app and it is not desirable to make changes to an app while a study is ongoing.

Participants also regularly had to be assisted during study participation, often to fix very small issues. It takes a team of a few people to stay on top of the entire study, especially during data collection. Being helpful, friendly, and patient with participants is important to keep them engaged and willing to document their lives each day. This is no different for traditional diaries that typically do not have the opportunity to realistically collect multiple weeks of participants’ daily lives. App development is more expensive up front, but the data are coded and ready to be analyzed as soon as they are collected, which is not the case when participants have to mail back paper-and-pencil diaries. All of these trade offs need to be considered when designing a similar framework.

It is our impression that TU+ is a state-of-the-art travel diary. Though we had hoped to say the same regarding time use diaries, our lacking timing information and other contextual information make TU+ incomplete compared to other studies purely dedicated to time use. Why so many participants chose to report activity "other" cuts our understanding of their full 24 hour day. At the same time, removing the option to label an activity as "other" frustrated users in earlier versions of the study and so we may have to keep it.

Furthermore, the collected data can be used to investigate a multitude of research questions. Transport researchers can begin to apply typical route and mode choice models and calculate policy-relevant time valuation metrics with these data. Social scientists can use the rich data to understand time use across different segments of the population, and how these change over time within and between individuals, especially in relation to travel.

The TU+ study is set to continue with an even larger sample of 30 days from at least 1,000 individuals collected over a period of six months. This longer sampling period will allow us to handle a manageable amount of communication that will inevitably be necessary between the participants and the research team and later to control for seasonality effects in mobility and time use patterns. Household members will be invited to participate in the study as well, which will be especially interesting for joint decision making, as other household members’ time use and mobility choices are tied to how a household shares its monetary resources as well, a dynamic well documented in Alho and colleagues (18) recent paper.
REFERENCES


