Data Collection Procedure

Before installing the application, we asked participants to read the information sheet and watch two YouTube tutorial videos explaining the application and the self-reports including four examples. The participants then installed the application from the Google Play Store. After opening the application for the first time, participants logged in with a username (i.e., an animal name) and password provided by the experimenter. After participants provided informed consent by selecting a checkbox, they were given a second chance to watch the tutorial videos. Next, the application requested participants to grant various device permissions.

In the final step, participants conducted a typing test on their default keyboard before the experiment and on the application keyboard before setting the application keyboard as the new default keyboard. The typing test consisted of six sentences in random order including two well-known pangrams (27, 30, 56, 37, 44, and 46 characters).

After the setup was completed, participants used their smartphones for 10 weeks in everyday life, filling in self-reports in regular intervals. We collected a total of 30,083 self-reports covering a large range of the valence-arousal-dominance space. Within the first week, we asked participants to fill in an online questionnaire on demographics and smartphone usage as well as the Patient Health Questionnaire and the Big Five Inventory 2 as measures of mental health and personality traits, respectively. At the end of the experiment, participants typed again the six sentences in random order on our keyboard and their default keyboard used before the experiment. Participants also completed an exit questionnaire on the self-reports regarding their level of understanding and the truthfulness and frequency of their responses. The exit questionnaire also probed their perception of the application's keyboard and smartphone usage. Finally, participants were asked to fill in the Patient Health Questionnaire and the Big Five Inventory for a second time.

Compensation

Participants were rewarded for their participation depending on their level of contribution and received between CHF 60 and CHF 120 for submitting an average of 3 and 6 self-reports per day, respectively. One participant was awarded an additional CHF 1000 from a lottery draw. Depending on the number of submitted self-reports, participants could reach three different levels providing a different number of tickets for the lottery: gold level (420 self-reports, 10 tickets), silver level (320 self-reports, 5 tickets), and bronze level (210 self-reports, 1 ticket). In addition, the participants with the highest number of average self-reports per day (after 70 days of participation) received additional tickets for the lottery (5 tickets for rank 1, 2 tickets for rank 2, and 1 ticket for rank 3).

Smartphone Usage

Participants used the smartphone in landscape mode in 0.96% of the sessions. To simplify the analyses, we excluded all sessions where the smartphone was used in landscape mode. Participants could pause the recording by enabling the private mode. Participants used private mode only 0.026% of the time.

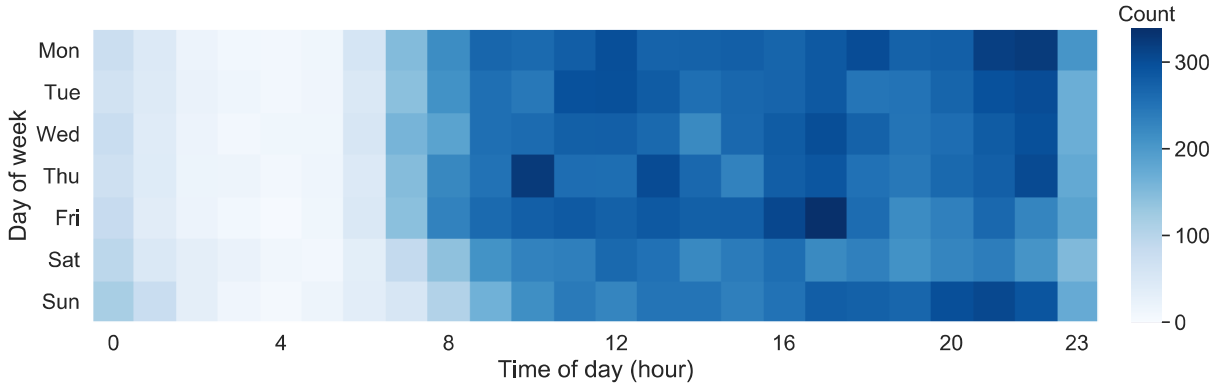

Figure 1: The distribution of the self-reports for the days of the week and the times of the day aggregated over all participants.

Self-Reports

Figure 1 shows the number of self-reports for the days of the week and the times of the day. Most self-reports were filled in between 7 a.m. and 11 p.m. Peaks are located on Mondays at 9 p.m. (318 self-reports) and 10 p.m. (323 self-reports), Thursdays at 10 a.m. (324 self-reports), and Fridays at 5 p.m. (338 self-reports).

A post-experiment questionnaire revealed that most participants completely (87%) or mostly (10%) understood the self-reports. All participants reported that they always (81%) or often (19%) filled in the self-reports truthfully. Most participants (59%) also felt that the self-reports had the right frequency. Only 11% and 30% of the participants reported that they found the self-reports being either too seldom or too often available, respectively.

Keyboard

In the modeling stage, we generated the heat maps from key pairs. We hypothesize that the typing speed of a key pair depends on the involved thumbs (e.g., key combinations typed with the left and right thumb might be faster). Key pairs can be located in the left or right part of the keyboard (left-left and right-right) or belong to both parts of the keyboard (left-right and right-left). We define the left (right) part of the keyboard as all keys to the left (right) of the vertical line going through the keys z, h, and v (the space bar is split in the middle). Here, we assume that the left and right thumb are used for the left and right parts of the keyboard, respectively. In a pre-experiment questionnaire, 84% of the participants reported to type with both hands (i.e., left and right thumb). An ANOVA revealed that there were significant differences in terms of the average typing speed for these four different key pair locations ($F(3, 336) = 110.395$, $p < 0.001$). Post hoc comparisons using the Tukey HSD test indicated that the mean typing speed for pairings of left-left (mean = 132 pixels per second (pps), SD = 61 pps), right-right (mean = 73 pps, SD = 35 pps), left-right (mean = 384 pps, SD = 211 pps) and right-left (mean = 505 pps, SD = 282 pps) were significantly different (all $p < 0.001$). Only the typing speed between left-left and right-right was not significantly different ($p = 0.135$).

At the beginning and at the end of the experiment participants typed six sentences in random order on their default keyboard and on our keyboard. A 2 (time) × 2 (keyboard) Aligned Ranked Transform (ART) ANOVA revealed that the typing speed was significantly higher ($F(1, 216) = 53.460$, $p < 0.001$) at the end of the experiment (mean = 1.380 characters per second (cps), SD = 0.255 cps) than at the beginning of the experiment (mean = 1.235 cps, SD = 0.240 cps). We also found that typing speed was significantly higher ($F(1, 216) = 5.571$, $p = 0.019$) for the default keyboard (mean = 1.325 cps, SD = 0.260 cps) than our keyboard (mean = 1.290 cps, SD = 0.235 cps). There was no interaction between keyboard and time ($F(1, 216) = 1.909$, $p = 0.169$). The higher typing speed at the end of the experiment might be because participants already saw the sentences at the beginning of the experiment. We conclude that although participants used diverse keyboards before the experiment, they quickly became familiar with our keyboard.