User-Donated Screenshots Analysis: Feasibility of a New Approach to Collect Objective Social Media App Usage in Adolescents

Yuning Liu[1] (✉), Geoff Klassen[2], Jenna Mee[2], Justin Pointer[2], Marvi Baloch[2], Laura Marciano[1], and Nathaniel Osgood[2]

[1] Harvard T.H. Chan School of Public Health, Cambridge, MA 02138, USA
{yuning_liu,lmarciano}@hsph.harvard.edu
[2] University of Saskatchewan, Saskatoon, SK S7N 5A2, Canada
{geoff.klassen,Jen917,justin.pointer,marvi.baloch,nathaniel.osgood}@usask.ca

Abstract. Objective data on social media use is now urgently needed for understanding its impact on adolescent well-being. Traditional objective social media data collection methods, such as data donation and passive sensing, face challenges including intrusiveness, privacy concerns, and limitations in adolescent—a critical demographic in this research area. In our study, we introduced a novel, less intrusive method using user-donated screenshots within an ecological momentary assessment (EMA) framework. We recruited 374 adolescents from Switzerland, who were instructed to capture and share three daily screenshots detailing their total and app-specific usage across screentime, activations, and notifications. From this, we collected 6,819 screenshots, with 25% of participants failing to submit any screenshots, 14% submitted incorrect or incomplete ones, while 64% provided complete data for more than five days. To process this data, we developed an image-to-text pipeline using Tesseract OCR that achieved a 96% average accuracy rate. This user-donated screenshot method proved to be less burdensome than traditional data donation, capable of capturing detailed app-specific usage across smartphone operating systems, and applicable among adolescents. Nonetheless, success of the user-donated screenshot approach hinges on user compliance. We analyze attrition sources and suggest six strategies to enhance future research, such as incentivizing participation, implementing pre-upload image checks, and improving participant onboarding and education.

Keywords: screenshot analysis · social media use · optical character recognition · ecological momentary assessment · adolescents

Y. Liu, G. Klassen—Co-first author.

R. Thomson et al. (Eds.): SBP-BRiMS 2024, LNCS 14972, pp. 80–89, 2024.
https://doi.org/10.1007/978-3-031-72241-7_8

1 Introduction

1.1 The Need for Objective Social Media Use Data

Social media and smartphone usage have surged over the last decade, coinciding with a marked decline in adolescent well-being [1–3]. This parallel trend has rung the alarms among researchers and policymakers. The U.S. Surgeon General's 2023 advisory on Social Media and Youth Mental Health emphasizes the need to closely examine social media behaviors to connect specific usage patterns with well-being outcomes [4]. While the link between social media use (SMU) and adolescent well-being is still contentious, understanding nuanced social media activities is central to this debate [5, 6].

The complexity of individual SMU patterns makes it challenging to assess their impact on well-being. Simply tracking total screentime is insufficient considering that identical usage durations can involve different apps and behaviors, leading to a variety of effects on well-being [7, 8]. For instance, two adolescents with equal screentime might engage differently—one frequently checking their phone, the other spending prolonged periods browsing. Such discrepancies highlight the limitations of self-reported data, which often underestimates actual usage or shows small correlation with objective usage data [9]. Therefore, objective data on SMU are now crucially needed for exploring the relationship between SMU and adolescent well-being, to inform the policy making agenda and develop public health interventions.

1.2 Limitations of Current Objective Social Media Use Data Collection Methods

Scholars have developed several methods to collect objective SMU data, known as digital trace or phenotyping. These methods capture user activities saved in online environments [10]. Objective SMU data can be collected via Application Programming Interfaces (APIs), app data donation, and passive sensing, though each method faces challenges [11]. API-based collection is limited by content restrictions, rate limits, and policy changes. Additionally, it is challenging to correlate SMU data obtained from APIs with well-being outcomes, which are typically assessed via surveys or professional diagnoses [11, 12]. App data donation packages consist of archives from social media platforms that each user can request to download. This approach is constrained by low compliance rates, delays in receiving data packages, and the complexity of extracting information from these packages, with different social media platforms presenting varied and complex data formats, including compliance with privacy issues [13].

Passive sensing apps generally only measure overall screen time and, except a few on Android, fail to track screen states by app due to iOS restrictions [14]. To address these challenges, the Human Screenome Project introduces Screenomic analysis, capturing smartphone screenshots every five seconds during active use [15, 16]. This approach represents a significant advancement in passive social media tracking, yet it raises concerns over its intrusiveness and the privacy risks [17]. Additionally, high-frequency data collection produces a vast amount of image data, which requires substantial computational resources to process and derive meaningful insights. In summary, effectively gathering granular and objective data on SMU, while adhering to international regulations and minimizing privacy concerns, poses a significant challenge.

1.3 The Present Study

This study introduces a novel method for collecting objective social media data through user-donated screenshots on app-specific usage, attempting to address several limitations of existing methods. This innovative approach utilizes an ecological momentary assessment (EMA) design, requesting participants to share screenshots of app usage from their smartphone Settings alongside survey responses in each EMA. This method ensures privacy protection and non-intrusiveness and is adaptable across various populations and settings. We developed an image-to-text pipeline to extract data from these screenshots. This paper presents our new approach, the user-donated screenshots analysis, and demonstrates the pipeline's performance through evaluation results. Additionally, we summarize the attrition patterns observed and discuss strategies to mitigate participant drop-off in future research. Ultimately, our study seeks to answer the following three main research questions: (i) How feasible is it to derive app-specific usage features from user-donated screenshots? (ii) How practical is it to collect such screenshots from participants in an EMA study? (iii) And how can future studies applying user-donated screenshot approach be designed to minimize attrition?

2 Method

2.1 Data Collection

The HappyB study studies social media use and well-being by collecting intensive longitudinal data from ecological momentary assessments (EMAs) among 374 Swiss adolescents (mean age = 15.71, SD = 0.82; 235 females, 62.8%) across 14 days in spring 2022. Participants were recruited from four high schools in Switzerland, encompassing students in the first and second high school grades, occurring between March 7th and March 23rd, 2022. Participants initially completed a baseline survey and engaged in EMAs through the Ethica app (now known as Avicenna), with assessments conducted three times daily at 12 pm, 6 pm, and 9 pm from Day 1 to Day 14. Self-reported questions on social media activity, experience, and well-being are included in each EMA. Furthermore, during the 12 pm EMAs, participants were directed to access their smartphone's Settings, capture three screenshots illustrating the total usage and the top three app usage statistics from the previous day—specifically, daily usage time, screen activations, and notification counts—and subsequently upload these screenshots to Ethica. A gamification system was developed inside the app to track participants' adherence. Each adolescent received a medal labeled in different ways (e.g., "conqueror", "hero") depending on their total submitted EMAs. At the end of the study, depending on adherence rates and then medal obtained, each participant received a gift card up to 20 CHF. Participants were then asked to upload the three screenshots to Ethica. Example screenshots are presented in Fig. 1. The study received the ethical approval by the IRB of USI, Università della Svizzera italiana, Lugano, Switzerland, and was supported by the Department of Education, Culture, and Sport of canton Ticino.

2.2 Image-to-Text Pipeline

We extract data on overall and app-specific usage for the three features—screentime, number of activations, and number of notifications—from the user-donated screenshots. Specifically, we extract the total use, identify the names of the top three most-used apps, and obtain the usage statistics for these apps from each screenshot. As shown in Fig. 1(C), the extracted information includes a total of 347 notifications received on the previous day, with WhatsApp accounting for the highest number at 267 notifications. TikTok and Instagram follow, with 24 and 17 notifications respectively.

Fig. 1. Example of user-donated screenshots on the screen time of App use (A), activation by App (B), and notification by App (C)

We developed an image-to-text pipeline using Tesseract Optical Character Recognition (OCR) to extract total usage and app-specific data from screenshots. The workflow of this pipeline is depicted in Fig. 2. To assess the accuracy of our text extraction, we randomly selected 600 screenshots, divided equally among iOS and Android devices for screen time, activation, and notification data, and annotated them to establish ground truth. For each screenshot, we generated seven metrics: total screentime/count, the app name with the highest screentime/count, screentime/count for this top app, names and screentime/counts for the second and third highest usage apps. The pipeline's performance was evaluated by comparing the accuracy rate—the proportion of screenshots where extracted text precisely matched the ground truth (Table 1).

2.3 Attrition Pattern Analysis

To evaluate the feasibility of using user-donated screenshots, we assessed not only the performance of the image-to-text pipeline but also analyzed the attrition patterns. This analysis aimed to identify how errors could arise from the user-donated screenshots.

Understanding the sources of these errors is crucial for assessing feasibility. We identified that issues occur at two levels: the user level (loss to follow-up or failure to submit screenshots) and the screenshot level (submission of incorrect or incomplete information). Consequently, we analyzed and summarized the sources of these problems at both levels, with the results depicted in the flowchart in Fig. 3.

3 Result

3.1 Image-to-Text Pipeline Achieves High Accuracy

Our image-to-text pipeline (Fig. 2) contains a series of steps to process screenshots from smartphones, utilizing Tesseract OCR for text recognition.

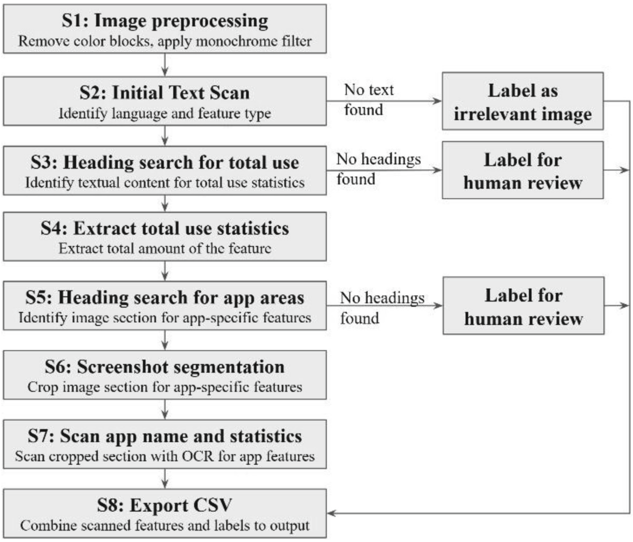

Fig. 2. Image-to-text pipeline in user-donated screenshot analysis in the HappyB study

First, it conducts image preprocessing by applying monochrome filters and removing color blocks. Second, the pipeline conducts a comprehensive scan of the textual content within the screenshot. From the textual content, the pipeline identifies the language, and the type of data displayed (screentime, activation, or notifications). It is worth noting that screenshots are submitted after participants self-identify their device OS (e.g. iOS or Android), so the pipeline does not include OS detection. Third, the pipeline conducts language-specific, type-based, and version-oriented searches for headings of total use statistics within the textual content from screenshots. The total use statistics are the total screentime, activations, and notifications from a day. For example, headings like "pickups" and "yesterday" are used to ascertain that the textual content after the heading is the total number of activations from yesterday. To reduce the errors from the OCR process and enhance the accuracy of heading detection, we implement an edit distance algorithm that matches the OCR-generated text against a pre-defined list of expected headings.

Fourth, with the identified headings for total use statistics, the pipeline extracts the total use statistics for screentime, activation, and notification from the given areas of the text. Fifth, the pipeline searches for headings for app-specific features. Headings like “most used” indicate that the textual content below is app-specific information. Edit distance-based search is also applied here. Sixth, with the identified headings for app-specific features, the pipeline segments the screenshot to isolate the specific areas containing app usage data. This segmentation is guided by spatial coordinates: the upper boundary is set by the heading for app-specific data (e.g., “most used”), the lower boundary is set by key text that appears below the app-specific data (e.g. “view more” or a subsequent heading), and the lateral boundaries are defined by the edges of the display or icons that appear on either side of the app data. Seventh, a second OCR process is then applied to segments of app-specific usage area to extract app names and associated usage statistics. To ensure the extracted app names are correctly aligned with their corresponding usage statistics, the pipeline conducts location-based logic checks to identify and address any discrepancies arising from the second OCR pass. Finally, the pipeline labels images that need human checkups, including the images with the absence of relevant headings, low OCR confidence scores, and identification of data as representing a weekly rather than daily overview. Data extracted from the images are systematically recorded in a CSV file for further analysis.

Table 1. The accuracy rate of the image-to-text pipeline in HappyB screenshot analysis

	IOS			Android		
	Screentime	Activation	Notification	Screentime	Activation	Notification
Total stats	0.93	0.91	0.99	0.95	0.98	0.99
APP #1 name	0.98	0.99	0.99	0.93	0.99	0.97
APP #1 stats	0.98	0.99	0.99	0.88	0.99	0.96
APP #2 name	0.96	0.99	0.99	0.92	0.99	0.95
APP #2 stats	0.97	0.99	0.97	0.93	1.00	0.96
APP #3 name	0.94	0.94	0.96	0.92	0.99	0.93
APP #3 stats	0.93	0.96	0.94	0.91	0.99	0.93

The accuracy rate of our pipeline, as presented in Table 1, achieved an average of 96.07% in extracting text across the seven metrics. Specifically, the pipeline demonstrated over 91% accuracy in identifying total usage statistics. For app-specific names, screentime, and activation/notification counts, it achieved an average accuracy rate of 96.11%. The notably high accuracy in the Android activation category is largely due to many screenshots from an older Android version in our dataset, which only records total activation counts without app-specific details.

3.2 Attrition Pattern at the User and Screenshot Level

The attrition pattern of the HappyB study is illustrated in Fig. 3. Initially, 374 participants were recruited for the baseline survey. On the subsequent day, 309 of these participants completed the first EMA. If each of the 309 participants had remained in the study and submitted screenshots for each of the 14 days, we would expect to receive 4,326 sets of screenshots covering screentime, activation, and notification. However, we received 2,273 sets of screenshots (equating to 6,819 screenshots) from 277.

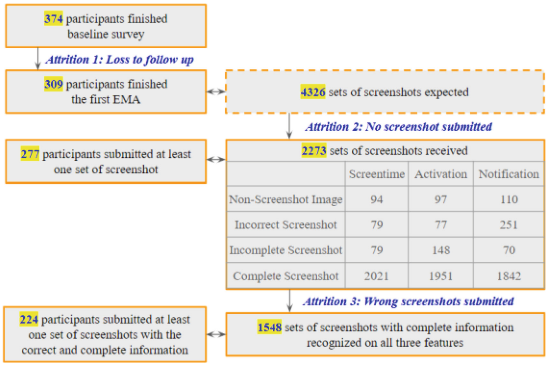

Fig. 3. Flowchart showing the attrition pattern in the HappyB EMA study with user-donated screenshots.

participants. Upon applying the image-to-text pipeline, we found that 301 submitted images (94 screentime, 97 activations, and 110 notifications) were not relevant screenshots but rather unrelated images such as personal photos. Additionally, 407 images (79 screentime, 77 activations, and 251 notifications) were screenshots but failed to include the total usage or the app-specific information we required, such as screenshots depicting only the battery state of the smartphone. Furthermore, 297 images (79 screentime, 148 activations, and 70 notifications) were incomplete, often due to participants cropping out key app-specific areas or usage details. Consequently, we obtained 1,548 sets of complete screenshots, covering all seven metrics across three screenshot types, with 224 participants providing complete data for all three features in at least one EMA.

We also visualize the user attrition pattern in screenshot donation in Fig. 4(A). On day 1, 277 participants submitted screenshots. From day 2 to 5, we observed a daily loss of 18 (6%) to 33 (12%) users. Between day 6 and 13, the attrition slowed to between 5 (2%) and 15 (5%) users per day. However, on the final day, we saw a significant drop with 60 users (21%) not submitting their screenshots. As shown in Fig. 4(B), among the 277 users who submitted screenshots, 59 (21%) submitted fewer than 3 sets, while 179 (65%) submitted more than 5 sets and 129 (47%) submitted more than 10 sets.

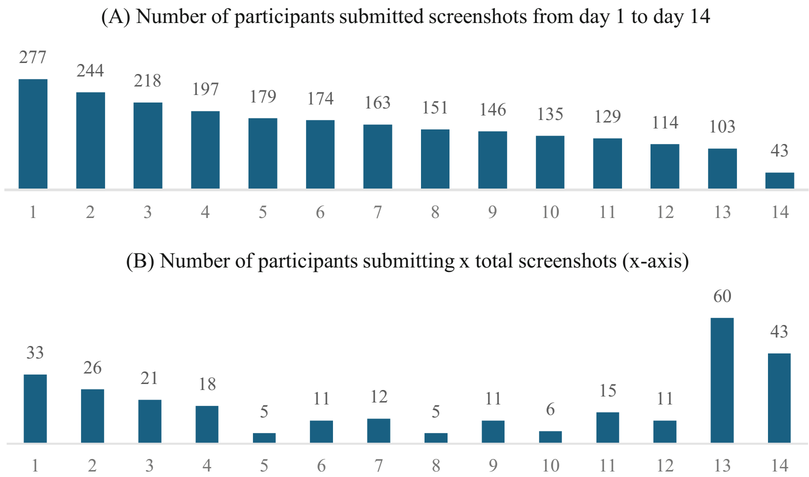

Fig. 4. Pattern of participants attrition in screenshot submission from day 1 to day 14 (A) and the distribution of the total number of screenshots submitted by participants (B)

4 Discussion

4.1 Summary of the User-Donated Screenshot Approach

In our study, we introduced a novel method for collecting social media data using user-donated screenshots. The user-donated screenshot method presents several advantages over traditional objective data collection methods. It is less intrusive than passive screen tracking apps and less burdensome for users than data donation packages. This approach allows for the collection of app-specific usage data across different smartphone operating systems. Notably, most passive sensing apps are limited to Android and cannot track on iOS. Furthermore, many data collection methods like API requests and some passive apps are restricted to adults, excluding adolescents—a key demographic for studying SMU and well-being. The user-donated screenshot approach allows for objective data collection in adolescents.

Success depended on two key factors: the accuracy of text extraction pipeline and user compliance. On the technical side, our image-to-text pipeline achieved an average accuracy rate of 96%, confirming the effectiveness of user-donated screenshots as a reliable source for app usage analysis. At the behavioral level, 26% of participants did not submit any screenshots, 14% submitted incorrect or incomplete ones, and 48% consistently provided complete screenshots for over five days. Our compliance rates align with other EMA study [18]. However, this approach has two main limitations. First, the user-donated screenshot method only captures basic metrics such as screen time, activations, and notifications, lacking depth in insights into social media activities and content. Second, its success heavily relies on user compliance. In the next section, we discuss and propose strategies to reduce attrition in future screenshot donation studies.

4.2 Solutions for Reducing Attrition in Future Screenshot Donation Studies

As far as we know, multiple research teams are now initiating plans for user-donated screenshot data collection. Therefore, it is crucial to summarize our learnings regarding the sources of attrition and strategies for reducing it in future studies. The taxonomy of attrition sources in user-donated screenshot analysis is detailed in Table 2. Of the 374 study participants, 277 submitted at least one set of screenshots, indicating a 26% non-submission rate. This may be attributed to practical burdens (participants feeling overwhelmed by the task), privacy concerns, and other barriers such as misunderstanding the screenshot donation process. To engage the 26% of participants who either lost to follow-up or did not submit any screenshots, future studies could: simplify the EMA design to lessen participant burden; enhance consent information and support to address privacy concerns; and incorporate gamification or rewards to incentivize participation.

Beyond cases of non-submission, 53 out of 374 participants (14%) submitted either non-screenshots or images lacking the requested information. This may result from accidentally uploading incorrect images or a lack of understanding of the required information. To address these issues, future studies could implement a built-in pre-check function in the EMA app that allows users to review images before submission. Additionally, integrating an asynchronous image-to-text pipeline that flags submissions lacking text could prompt users to confirm their uploads. Enhancing education during participant onboarding can also ensure a clearer understanding of the study's objectives.

Table 2. The taxonomy of the source of attrition in user-donated screenshot analysis

<table>
<tr><th>Source of attrition</th><th>Solution in future studies</th></tr>
<tr><td>1. Loss to follow-up from EMA
Due to practical burden, privacy concerns, or barriers</td><td rowspan="2">a. Simplify EMA design
b. Address privacy concerns
c. Incentivize participants</td></tr>
<tr><td>2. Stayed in EMA but no screenshots submitted
Due to privacy concerns, barriers, or knowledge gap</td></tr>
<tr><td>3. Non-screenshot images submitted
Due to uploading the wrong image from album</td><td rowspan="3">d. Implement screenshot pre-check
e. Async verification via OCR
f. Enhance participant onboarding and education</td></tr>
<tr><td>4. Screenshots lack requested information
Due to lack of understanding of the study instruction</td></tr>
<tr><td>5. Screenshots with incomplete information
Due to lack of understanding of the study instruction</td></tr>
</table>

Acknowledgments. This study was funded by the Swiss National Science Foundation (grant P500PS_202974) and the NIH/NIMH (grant 1R21HD115354-01).

References

1. Massarat, E.A.V.: Risa Gelles-Watnick and Navid: Teens, Social Media and Technology (2022)

2. NW, 1615 L. St, Washington, S. 800, Inquiries, D. 20036 U.-419-4300 | M.-857-8562 | F.-419-4372 | M.: Social Media Fact Sheet (2024). https://www.pewresearch.org/internet/fact-sheet/social-media/
3. CDC: Youth Risk Behavior Surveillance Data Summary and Trends Report: 2011–2021 (2021). https://www.cdc.gov/healthyyouth/data/yrbs/pdf/YRBS_Data-Summary-Trends_Report2023_508.pdf
4. Advisory of U.S. Surgeon General: Social Media and Youth Mental Health — Current Priorities of the U.S. Surgeon General. https://www.hhs.gov/surgeongeneral/priorities/youth-mental-health/social-media/index.html
5. High, A.C., Ruppel, E.K., McEwan, B., Caughlin, J.P.: Computer-mediated communication and well-being in the age of social media: a systematic review. J. Soc. Pers. Relationsh. **40**, 420–458 (2023). https://doi.org/10.1177/02654075221106449
6. Meier, A., Reinecke, L.: Computer-mediated communication, social media, and men-tal health: a conceptual and empirical meta-review. Commun. Res. **48**, 1182–1209 (2021). https://doi.org/10.1177/0093650220958224
7. Kross, E., Verduyn, P., Sheppes, G., Costello, C.K., Jonides, J., Ybarra, O.: Social media and well-being: pitfalls, progress, and next steps. Trends Cogn. Sci. **25**, 55–66 (2021). https://doi.org/10.1016/j.tics.2020.10.005
8. Orben, A.: Teenagers, screens and social media: a narrative review of reviews and key studies. Soc. Psychiatry Psychiatr. Epidemiol. **55**, 407–414 (2020). https://doi.org/10.1007/s00127-019-01825-4
9. Ohme, J., Araujo, T., de Vreese, C.H., Piotrowski, J.T.: Mobile data donations: as-sessing self-report accuracy and sample biases with the iOS Screen Time function. Mob. Media Commun. **9**, 293–313 (2021). https://doi.org/10.1177/2050157920959106
10. Sultan, M., Scholz, C., van den Bos, W.: Leaving traces behind: using social media digital trace data to study adolescent wellbeing. Comput. Hum. Behav. Rep. **10**, 100281 (2023). https://doi.org/10.1016/j.chbr.2023.100281
11. Ohme, J., et al.: Digital trace data collection for social media effects research: APIs, data dona-tion, and (Screen) tracking. Commun. Methods Meas. 1–18 (2023). https://doi.org/10.1080/19312458.2023.2181319
12. Davidson, B.I., et al.: Platform-controlled social media APIs threaten Open Science (2023). https://osf.io/ps32z
13. van Driel, I.I., Giachanou, A., Pouwels, J.L., Boeschoten, L., Beyens, I., Valkenburg, P.M.: Promises and pitfalls of social media data donations. Commun. Methods Meas. **16**, 266–282 (2022). https://doi.org/10.1080/19312458.2022.2109608
14. Stier, S., Breuer, J., Siegers, P., Thorson, K.: Integrating survey data and digital trace data: key issues in developing an emerging field. Soc. Sci. Comput. Rev. **38**, 503–516 (2020). https://doi.org/10.1177/0894439319843669
15. Reeves, B., Robinson, T., Ram, N.: Time for the human screenome project. Nature **577**, 314–317 (2020). https://doi.org/10.1038/d41586-020-00032-5
16. Reeves, B., et al.: Screenomics: a framework to capture and analyze personal life experiences and the ways that technology shapes them. Hum. Comput. Interact. **36**, 150–201 (2021). https://doi.org/10.1080/07370024.2019.1578652
17. Ram, N., Haber, N., Robinson, T.N., Reeves, B.: Binding the person-specific approach to modern AI in the human screenome project: moving past generalizability to transferability. Multivar. Behav. Res. 1–9 (2023). https://doi.org/10.1080/00273171.2023.2229305
18. Tonkin, S., et al.: Evaluating declines in compliance with ecological momentary assessment in longitu-dinal health behavior research: analyses from a clinical trial. J. Med. Internet Res. **25**, e43826 (2023). https://doi.org/10.2196/43826