The Idiosyncrasies of Everyday Digital Lives: Using the Human Screenome Project to Study User Behavior on Smartphones

Miriam Brinberg[a], **Nilam Ram**[a], **Xiao Yang**[a], **Mu-Jung Cho**[b], **S. Shyam Sundar**[a], **Thomas N. Robinson**[b], **Byron Reeves**[b]

[a]Pennsylvania State University

[b]Stanford University

Abstract

Most methods used to make theory-relevant observations of technology use rely on self-report or application logging data where individuals' digital experiences are purposively summarized into aggregates meant to describe how the average individual engages with broadly defined segments of content. This aggregation and averaging masks heterogeneity in how and when individuals actually engage with their technology. In this study, we use screenshots ($N > 6$ million) collected every five seconds that were sequenced and processed using text and image extraction tools into content-, context-, and temporally-informative "screenomes" from 132 smartphone users over several weeks to examine individuals' digital experiences. Analyses of screenomes highlight extreme between-person and within-person heterogeneity in how individuals switch among and titrate their engagement with different content. Our simple quantifications of textual and graphical content and flow throughout the day illustrate the value screenomes have for the study of individuals' smartphone use and the cognitive and psychological processes that drive use. We demonstrate how temporal, textual, graphical, and topical features of people's smartphone screens can lay the foundation for expanding the Human Screenome Project with full-scale mining that will inform researchers' knowledge of digital life.

Declarations of interest: none.

1. Introduction

Individuals' digital experiences are now characterized by short segments of exposure and use and many quick switches into and among radically different kinds of content (Yeykelis, Cummings, & Reeves, 2014; Reeves et al., 2019). Most psychological theories about technology use, and most of the methods used to make theory-relevant observations, rely on self-report or application logging data where individuals' digital experiences are purposively summarized into aggregates meant to describe how the average individual (typically, a young adult) engages with very broadly defined segments of content (e.g., social media, news, videos). This aggregation and averaging causes two problems. First, it masks and muddles the heterogeneity in how (structure), what (content), why (function), where (context) and when (time) individuals actually engage with their devices. Second, it diminishes the importance of explaining why heterogeneity occurs and why it must be considered as more than mere measurement error in theories about digital life. While aggregation and averaging individuals' experiences are not the only limitations in current studies of media use, we highlight these issues because they currently hinder theoretical and applied work. Specifically, researchers and those working in industry may not have the data needed to develop precise within-individual predictions, and are thus relying on imprecise knowledge about the "average" person or the "average" day. Furthermore, the current focus on describing broad, non-overlapping categories of web content and applications (e.g., social media, news, videos) does not provided detailed knowledge about homogeneity and/or heterogeneity of the specific content that users engage within and across the many platforms they access each day. Rather than analyzing the data using typical and often-used general (and commercially-based) media categories (e.g., social media, news), we extract features from screens and describe these features in dimensional and topic-oriented ways. For example, we demonstrate identification of content pertinent to specific topics (e.g., nutrition, purchasing) that can span multiple commercially-based media categories. The shift towards dimensional specificity adds new detail to our knowledge of individuals' actual device use.

In this paper, we use the Screenomics approach (Reeves et al., 2019) to address current limitations in describing the heterogeneity of individuals' experiences and how these experiences manifest across content areas. We begin by describing some current methods for describing daily digital life (e.g., URL and application logging tools). We then forward use of the *screenome* – screenshots recorded every five seconds a smartphone is activated – as a new method to observe and study the full content and flow of information on smartphone screens, and demonstrate how temporal, textual, graphical, and topical features of 132 individuals' smartphone screens can provide detailed insight into individuals' everyday digital experiences. Following presentation of results that highlight the substantial between-person and within-person heterogeneity in how individuals switch among and titrate their engagement with different content, particularly in how textual and visual content flows change throughout the course of the day, we conclude with a discussion of the results,

privacy concerns related to Screenomics, and current limitations and future directions of Screenomics.

Generally, one of the foremost goals of Screenomics is accurately describing daily digital life. With more accurate and detailed descriptions of everything individuals see and do on their screens, researchers can test and potentially improve upon current theories that address the underlying mechanisms and effects of use. As well, seeing and studying digital media use at the resolution of the screenome will propel development of new theory. In turn, improved understanding can lead to better prediction of media use behaviors and development of novel and precise interventions (e.g., mHealth) that help individuals benefit from their screen use and limit adverse impacts through personalized reorganization of content (e.g., changing the flow of content based upon individuals' use patterns). Our goals for this paper are to (1) demonstrate the use of Screenomics as an approach that is suitable to address a number of research questions about media use, (2) present exemplar uses of the Screenomics approach and how the resulting findings inform current media theories, and (3) lay the foundation for expanding the Human Screenome Project and doing full-scale mining that will inform many researchers' and technology designers' knowledge of digital life.

2.   Background

2.1   In-situ Study of Digital Life

Researchers often use data from surveys and questionnaires to understand people's smartphone use. Given the difficulty participants have trying to report accurately on their quickswitching, highly frequent behaviors, inferences of actual use from self-report measures are severely limited and rather unreliable (Andrews, Ellis, Shaw, & Piwek, 2015). To address the limited reliability of self-report measures, and to better capture people's actual device use, researchers have developed a variety of URL and application logging tools (Cockburn & McKenzie, 2001; Jansen & Spink, 2006; Tossell, Kortum, Rahmati, Shepard, & Zhong, 2012; White & Huang, 2010). These measures, however, also have some limitations. While URL and application logging methods obtain more accurate information about when and what websites people visit and what applications people use, the available tools are often not able to describe (1) the features of the media stimuli that individuals encounter, (2) the specific content and diversity of activities that can occur within a particular application, and (3) the overlap of content across platforms and/or time. The inability to describe the granularities of content and flow of information could be the reason why previous research has found no systematic relationship between psychologically meaningful constructs and technology use, such as the associations between psychological well-being and digital media exposure (Hancock, Liu, French, Luo, & Mieczkowski, 2019). New tools are needed for observing digital life, particularly tools that facilitate observation of the detail and fine granularity of digital content and actions that people actually consume, share, and experience.

2.2   Screenomics

The smartphone screen is a venue for consuming, sharing, experiencing, and, from a scientific perspective, observing the entirety of digital life (Reeves, Robinson, & Ram,

2020). This paper makes use of new methods for obtaining the intensive longitudinal data streams needed for comprehensive description of the ephemeral, cross-platform, and cross-application digital environments that individuals encounter during *in-situ* engagement with search results, ads, news, games, articles/stories, videos, social relationships, and whatever else they do on their phone.

In brief, Screenomics is made possible by a new data collection and analysis framework for capturing, storing, visualizing, and analyzing the full record of what individuals view and interact with on smartphone and laptop screens over time – the *screenome* (Reeves et al., 2019; Ram et al., 2020). The framework includes software and analytics that collect screenshots from smartphone screens every few seconds when the devices are on, extract text and images from the screenshots, and automatically categorize and store screen information with associated metadata. The system architecture is designed so that the entire collection and text and image extraction pipeline can run automatically (with some small amount of maintenance and checking). Using these extracted features, researchers are then able to create visualizations and conduct analyses using the visualization or analysis software program(s) of their choice. By collecting screenshots every few seconds and sequencing those screenshots into a screenome, researchers obtain a new "microscope" for studying how, when, and for what purposes individuals actually use their phones and screens as they go through everyday life. While Reeves and colleagues (2019) presented the overarching Screenomics framework, this paper provides the first systematic description of the screenome and an analysis pipeline for examining screenome data at scale.

The temporal density of the data (screenshots taken at five-second intervals whenever the smartphone screen is activated) and the coverage of the data (all the pixels on the entire screen) provide a detailed account of the entire stimuli (i.e., content) that is presented to users on their smartphone screen – and thus, overcomes some of the limitations of self-report or other logging data about smartphone use. In the same manner that real-time recordings obtained in laboratory-based studies provide for detailed study of users' experience with specific applications or digital environments (Kaufman et al., 2003; Sundar, Bellur, Oh, Xu, & Jia, 2013), the screenome provides a full record of user experience, but *in-situ*, and at the speed at which behaviors of interest may occur (seconds time-scale).

Using the Screenomics framework, the present study provides a first examination of how 132 smartphone users engage with all the media content that appears on their screen. In-situ media experiences—which are fast, fluctuating rapidly, and hard to sample—are known to trigger a variety of psychological effects and/or reflect users' current psychological state. Using this new data stream, we extract a first set of metrics that parse, quantify, and describe psychologically meaningful aspects of individuals' everyday digital experiences.

2.3   Textual, Topical, Graphical, and Temporal Content/Behavior

Content is a fundamental descriptor of individuals' digital lives. Research about media behavior and effects aims to identify pairs of stimuli and responses that describe how individuals experience media. In studying media behavior and effects, psychologists and media researchers seek to characterize media content according to some stimuli of theoretical interest (Reeves, Yeykelis, & Cummings, 2015). Describing real-world media

experience has always been a challenge, however, as conventional experimental and experience sampling methods have limitations. For experimental studies, creating a sample of stimuli that matches what is experienced in the real world is a prominent methodological issue (Judd, Westfall, & Kenny, 2012; Reeves et al., 2015). Studies that rely on experience sampling methods are constrained by obtrusiveness and sampling frequency and interval, often capped at 8-12 observation points per day for sampling periods longer than 2-3 weeks (Christensen, Barrett, Bliss-Moreau, Lebo, & Christensen, 2003; Delespaul, 1992; Reis & Gable, 2000). Many important features of media content can hardly be replicated in a controlled setting as they are actually experienced in the real world and require much more intensive sampling.

Screenomics solves these issues by simply capturing and analyzing screenshots of all the stimuli that appear on users' smartphone screens as they go about their daily lives; thus, giving researchers the *in-situ* observations needed to understand the uses and gratifications of modern technology (Sundar & Limperos, 2013). We describe potential uses of Screenomics along with the examples used in this paper with the hope that these uses spark readers' development of additional ideas. For instance, researchers may be interested in studying media multitasking. Even when researchers use application logging tools to study media multitasking in individuals' daily lives, these tools miss task-relevant content that may be present across platforms (e.g., working in a Word document on a research paper, then switching to e-mail to see what a coauthor said about the project). In contrast, Screenomics captures, moment-by-moment and in fine detail, all the content on individuals' screens and how they switch among that content. In this paper, we use newly available Screenomics data to describe when and for how long individuals are exposed to the stimuli on their screens, chronicle how some general characteristics of the textual and graphical stimuli appearing on the screen change across time (within-day), and illustrate how specific domains and topics of content (e.g., nutrition, public affairs, purchasing) are spread throughout individuals' everyday smartphone use. We note, however, that our measures extracted from the screenome do not capture all textual and graphical stimuli with complete precision; we discuss these imprecisions in sections 3.3 Measures and 5.2 Limitations and Future Directions.

Generally, the text present on individuals' screens can inform researchers about a variety of behaviors or constructs of interest, such as potential cognitive load the user experiences or the topics to which the user is exposed. To illustrate how Screenomics data can be used to examine text, we examine word count, word velocity, and the presence of three exemplar topics of interest (nutrition, public affairs, purchasing). *Word count* is a time-varying textual feature of media content. Word count, one potential measure of cognitive load, has previously been linked to social media engagement (Shami, Muller, Pal, Masli, & Geyer, 2015), expression of emotional distress on social media (Brubaker, Kivran-Swaine, Taber, & Hayes, 2012), and is said to be a proxy for amount of communication (Tausczik & Pennebaker, 2009).

Screenomics data also allows us to examine when individuals engage with specific types of content, in such popular domains of media consumption as health, politics, and e-commerce. Collecting data at a fast time-scale (i.e., seconds) allows us to study how exposure to

particular kinds of stimuli are distributed across individuals' actual digital environment. We can now see when particular threads of content appear and resurface, across applications and across time. In this paper, we use three topics (nutrition, public affairs, purchasing) as exemplar illustrations of how Screenomics may inform research in a wide variety of fields, and eventually support development of personalized and precise, real-time interventions and design of consumer products in the fields of health, politics, and e-commerce. For instance, screenome-based analysis of nutrition-related content can provide researchers with better quantifications of the actual nutrition-related images and text appearing on screens and how often individual users view, create, and share it, which can then inform understanding of how food exposure via media is associated with and used to change actual food consumption. Other methods are not able to provide even the basic information about media exposure because they do not observe the actual screens viewed. Similarly, analysis of public affairs content provides a measure of news exposure across platforms that can inform not only when people access "news" via specific news-related applications, but can also identify when news-related topics are seen on non-traditional news sites (e.g., on social media). Analysis of payment-related content provides more precise, cross-platform knowledge about when people are shopping and provides previously unseen detail about what people do before and after purchase decisions. In sum, we chose nutrition, public affairs, and purchasing topics as illustrative examples of how researchers and designers can use the text extracted from screenshots to examine a particular topic of interest across application categories.

In addition to facilitating analysis of text-based stimuli, Screenomics data also facilitate analysis of graphical stimuli. For instance, researchers can use computer vision to recognize whether particular images (e.g., a company's logo) are present in a given screenshot, and thus determine if, when, and in what context users are exposed to specific kinds of content (e.g., advertisements). In our illustration, we focus on general summaries of the graphical features of Screenomics data, specifically image complexity. *Image complexity*, like word count, is a time-varying graphical feature of media content that indicates extent of visual stimulation. The complexity of visual information is closely associated with the depletion of cognitive capacity (Thorson, Reeves, & Schleuder, 1985) and has a significant impact on recognition memory (Reeves, Thorson, & Schleuder, 1986).

While the absolute values of *word count* and *image complexity* provide relative scales for comparing different media stimuli, the associated velocity metrics—the rates of change in text and graphics—also have important psychological implications. The velocity of media experience has been linked to emotion outcomes. For example, video games are found to be most enjoyable when the speed of the game is matched with players' mental capacity (Sherry, 2004). The velocity of scene progression in films also changes individuals' recognition of film genres (Visch & Tan, 2007). Following these lines of research, designers of media for smartphone screens may be interested in how quickly people move through content on their screen (i.e., scrolling behaviors). Given that interactivity is indicative of cognitive function, psychologists may also be interested in tracking how individuals are flowing through textual or graphical content, particularly in relation to arousal levels (e.g., sleep-wake cycles).

The textual and graphical features described above take advantage of the granularity afforded by Screenomics data. The fine-grained sampling of Screenomics data also allow for the examination of temporal features of smartphone use and behavioral phenomena at multiple time-scales. For instance, Screenomics data can be used to examine "short" time-scales, such as the quick-switching behaviors that occur during each individual smartphone use session (i.e., from phone on to phone off). Screenomics data, however, can also be aggregated to examine "longer" time-scales of interest, such as hours, days, and seasons. In our illustrative example, we examine how certain behaviors and stimuli manifest at the session time-scale (e.g., how long does a typical smartphone on-off period last?), the hourly time-scale (e.g., how do textual and graphical features change from hour-to-hour throughout the day?), and the daily time-scale (e.g., how long do individuals use their smartphone on an average day?).

In sum, the collection of screenshot data at a fine-grained time-scale for an extended period of time (e.g., days, months) allows researchers to examine individuals' digital media experiences at the time-scale at which they occur (anywhere from seconds to weeks), and to parse and describe in detail how the contents of those experiences change over time.

2.4   The Present Study

Most current research on smartphone use has relied on self-reports and URL/application logging tools. We build upon this foundational research using Screenomics, and begin describing and examining features of the actual stimuli and content that users engage with on their smartphone screens during everyday life. Specifically, the goal of this study is to describe the timing and patterning of smartphone content using an illustrative set of temporal, textual, graphical, and topical information extracted from the complete record of adults' smartphone activities. Our findings describe new details about the idiosyncratic nature of individuals' daily digital life and lay the foundation for extending the Human Screenome Project for the study of a wide variety of behavioral phenomena.

3.   Method

3.1   Participants

Participants were 132 adult smartphone users recruited from focus groups on media use conducted in New York City, Chicago, and Los Angeles using a convenience sampling method. Users were eligible for participation in the focus group if they had an Android smartphone and a laptop (either Windows or Mac). At the completion of the focus group session, the 178 participants were approached about participating in a study on mobile device use. Of the 86% ($N = 153$) who enrolled in our study on mobile device use, 132 users successfully completed at least two days of the study. These $N = 132$ (64 female, 68 male) individuals were between age 18 and 50 years ($M_{Age} = 34.67$, $SD_{Age} = 8.72$) and self-identified as Caucasian ($n = 75$), African American/Black ($n = 18$), Asian American ($n = 11$), or Latinx ($n = 28$). Most were single or living with a significant other ($n = 89$), and some were married ($n = 37$) or divorced ($n = 6$). Income ranged from '<$35,000' ($n = 17$) to '>$100,000' ($n = 17$), with a median income of '$50,000 – $74,999' ($n = 45$; $n = 21$ with income of '$35,000 - $49,999'; $n = 32$ with income of '$75,000 - $99,999').

3.2   Procedure

After expressing interest in the study, participants were given additional information about (1) the goals of the study (understanding smartphone use in daily life), (2) what would be collected from their devices (screen content via screen captures; GPS information – not used in the current study, but has been used in other projects examining the intersection of context and media use, Roehrick, Brinberg, Ram, & Reeves, 2019), and (3) the potential risks and benefits of their participation in the study (risks to privacy similar to those in everyday digital device use; no direct benefits other than monetary compensation). Those who agreed to participate and provided informed consent for data collection were asked to use their devices as usual during the next weeks while software (i.e., an application that was created specifically for the Human Screenome Project that is available for participants on the Google Play Store) loaded on their devices unobtrusively collected screenshots of whatever appeared on their device's screen every five seconds while in use. Over the next few weeks, screenshots were collected every five seconds when the screen was on using a rigorous privacy-preserving protocol. Specifically, at periodic intervals that accommodated constraints in bandwidth and device memory, bundles of screenshots were automatically encrypted and transmitted to a secure research server only accessible to the research team. Additional measures to ensure participant privacy are described in section 5.1 Privacy Considerations. At the end of the collection period, the software was removed from the devices, and the participants were thanked and given their compensation ($100 for the focus group, $200 for 10 days of Screenomics data collection, and an additional $100 for one month of Screenomics data collection).

Prior to data collection, we engaged in multiple rounds of pilot testing and human labeling of screenshots to ensure that our Screenomics application obtained accurate and high-quality screenshots of phone screen content across multiple types of Android devices and OS versions. The application does indeed capture the screens accurately and at high resolution. The quality of specific features subsequently extracted from the screenshots is assessed using samples of human-coded data and described in greater depth below.

In the current data set, a subset of users' laptop use was also observed using Screenomics; however, those data introduce additional computational complexities (e.g., identification of the primary window) and are not reported in this manuscript. This paper reports only on the screenshots obtained from individuals' smartphones. In total, we observed over 8,000 hours of in-situ smartphone use during 2,791 person-days of data collection from 132 adults (total of 6,002,552 screenshots). These screenshot sequences – screenomes – offer a rich record of how and when these individuals engaged with digital media on their smartphones during their everyday lives.

3.3   Measures

Screenomes were prepared for analysis using a data processing pipeline that extracted a variety of temporal, textual, graphical, and topical features from each screenshot. The "inputs" of the data processing pipeline are the time-stamped screenshots and GPS information collected from users' phones. From these pixel arrays and the time and location vectors, we extracted the selected features as "outputs" (e.g., image velocity). While these

extracted features contain some imprecision, these inaccuracies are treated as measurement error in the analyses. Description of the larger Screenomics framework, data collection efforts, and processing architecture can be found in Reeves et al (2019). Details relevant to this study and analysis are given below.

3.3.1   Temporal Features—Contiguous sequences of time-stamped screenshots obtained every five seconds that the device was activated allow for fine-grained description and analysis of the temporal organization of an individual's smartphone use. Temporal features were examined in two ways to benchmark this study to other application logging studies (e.g., Andrews et al., 2015). First, we simply inferred *total use time* from the total number of screenshots obtained per day. Second, we defined a smartphone *session* as the length of time the smartphone screen was on continuously. Each time the screen was turned off for ten or more seconds, the next activation marked the beginning of a new session (van Berkel et al., 2016).

3.3.2   Textual Features—Text was extracted from each screenshot using a customized Optical Character Recognition module (Chiatti et al., 2017; accuracy of 74%). In brief, each screenshot was (1) converted from RGB to grayscale to facilitate separation of foreground text from backgrounds that may be different colors, as well as to provide a fixed number of color categories (256 grayscale) for calculation of entropy (described below), (2) binarized to discriminate foreground from surrounding background, (3) segmented into blocks of text and images using dilation, and (4) passed through a Tesseract-based word/character recognition engine. The resulting ensemble of text obtained from each screenshot was treated as a bag-of-words and quantified in various ways.

Word count is a simple description of the quantity of text that an individual is exposed to on their smartphone screen and interpreted as a measure of cognitive load. We quantified word count for each screenshot as the number of "words," defined specifically as any string of characters separated by a space, that were extracted using our OCR module. Examination of the distribution of scores suggested that word count was inflated (e.g., 500+ words on a screen) for some screenshots that were, for instance, taken and read in a horizontal rather than vertical orientation. These values were trimmed by recoding all values > 150 as 150. Because of this trimming, and some other systematic inaccuracies that occur when using OCR extraction on screenshot images (e.g., battery icon being read as an E, skipping foreign language words), word count is interpreted as relative indicator of textual information on a screen, rather than the absolute amount of meaningful text that was on the screen.

Word velocity is a description of how quickly the textual content is moving on the smartphone screen (e.g., when scrolling through text). We quantified word velocity as the difference between bag-of-word ensembles obtained from two temporally consecutive screenshots, specifically, the number of unique new words that did not appear in the screenshot five seconds earlier. Similar to *word count, word velocity* was trimmed to range from 0 to 150.

3.3.3   Graphical Features—Graphical features of each screenshot were quantified using parallel conceptions of quantity and velocity. *Image complexity*, one way of operationalizing

visual stimulation, was quantified for each screenshot as the across-pixel heterogeneity (i.e., "texture") of a screenshot, specifically calculated as Shannon's entropy across the 256 gray-scale colors (Shannon, 1948). Image complexity ranged from 0 to 8, with higher numbers indicating greater image complexity.

Image velocity, the first derivative of the image complexity time-series, is a description of how quickly the graphical content is moving on the smartphone screen. Conceived as an indicator of visual flow (Wang, Lang, & Busemeyer, 2011), image velocity was quantified for each screenshot as the absolute value of the change in image complexity from the temporally prior screenshot.

3.3.4   Topical Features—Together, the textual and graphical features provide quantitative descriptions of each screenshot that are indicative of general levels of exposure, extent of cognitive processing, or amount of psychological stimulation. They do not provide any information about the types of content that appeared in the screenshot sequence. To illustrate how screenomes can be used to study specific kinds of content, we make use of two more features that are easily extracted from our current data processing pipeline.

Similar to other logging paradigms, information about what *application* was present on the screen at any given time can be extracted from the screenome. Here, these commercially-oriented labels were extracted using a machine learning algorithm that was trained on a subset of labeled data. In brief, trained coders manually labeled a subset of 87,182 screenshots with the name of the specific applications appearing on the screen (done using Datavyu; Gilmore, 2016). This set of labeled data was then used to train a machine learning algorithm that could accurately identify the application based on 11 textual and graphical features that had already been extracted from each screenshot (e.g., word count, word velocity, image complexity, image complexity velocity, session length). After tuning, we obtained a random forest (Hastie, Tibshirani, & Friedman, 2009; Liaw & Wiener, 2002) with 600 trees that had an out-of-bag error rate of 15%, and used that model to propagate application labels to the rest of the data. Parsimonious descriptions of individuals' use of 315 unique applications were obtained by categorizing applications by type. Following previous research (Böhmer, Hecht, Schöning, Krüger, & Bauer, 2011; Murnane et al., 2016), applications were grouped into higher-order commercial categories based on the developer-specified categories used in the official Android Google Play marketplace. Thus, each screenshot was labeled with both the specific *application name* (e.g., Snapchat, Clock, Instagram, YouTube) and 1 of 26 general *application types*: Books & References, Business, Communication, Education, Entertainment, Finance, Food & Drink, Games, Health & Fitness, House & Home, Lifestyle, Maps & Navigation, Music & Audio, News & Magazines, Personalization, Photography, Productivity, Puzzle, Shopping, Social, Sports, Study (our data collection software), Tools, Travel & Local, Video Players & Editors, and Weather.

Complementary to the representation of smartphone content in terms of commercially oriented distinctions among applications, we also labeled each screenshot with respect to whether the actual text appearing on the screen was related to three illustrative topics: *nutrition, public affairs*, and *purchasing*. For each topic, we developed a dictionary that

contained a non-exhaustive list of words that would indicate that the material on the screen was related to the target topic. The dictionaries were created by graduate students who specialize in these areas, drawing from existing domain dictionaries and their expert knowledge of media content, for the purpose of querying the screenome data to study the potential effects of exposure to information on these topics. Our intent here was to be illustrative rather than comprehensive. The nutrition dictionary contained 61 words and phrases and included both formal food-related language used in nutrition literature (e.g., "carbohydrate") and informal food-related language used in social media (e.g., "hangry"). The public affairs dictionary contained 83 words and phrases and included words relevant to news and politics at the time of data collection (the United States in 2017), including words such as "border" and "health care." The purchasing behaviors dictionary contained 19 words and phrases related to shopping online, such as "shipping," "shopping cart," and "billing address." The bag-of-words obtained from each screenshot was then compared to each dictionary using standard dictionary look-up. Screenshots that contained one or more of the words in the dictionary were coded as containing that topic. In the present study, we do not examine the simultaneous occurrence of multiple topics on a single screenshot (e.g., nutrition and purchasing words both present on a screenshot). At the general level, the evaluation of topics and the intersection of terms across topics will depend on how researchers operationalize specific topics in relation to their underlying research area and question. In our case, the selected terms did not overlap across dictionaries, so the three illustrative topics were "orthogonal" in the text. However, we can imagine instances when overlap may occur (e.g., dictionaries for "depression" and "suicide"). Noting that text- and image-based dictionaries can be developed for any topic of interest, we use this analysis to illustrate how screenomes can be used to study individuals' engagement with specific topics both across time and across applications. We also acknowledge the non-exhaustive nature of our dictionaries and note that the imprecision of the extracted text may result in under-estimation of topic identification.

4. Results

Features were extracted from all 6,002,552 of the screenshots that comprise the 132 screenomes that ranged in length from 2 to 38 days (M = 21.14, SD = 10.91, Median = 22.50). In the sections below, we use these screenomes to describe temporal, textual, graphical, and topical aspects of individuals' everyday smartphone use.

4.1 Temporal Aspects of Smartphone Use

Individuals' smartphone use is non-continuous, fragmented, and scattered throughout the day in irregular and idiosyncratic ways. Figure 1 depicts week-long snippets of screenomes obtained from three illustrative participants. Each row of colored vertical bars indicates when the participant's screen was activated during each five-second interval of a 24-hour period, with the colors and black shapes indicating the type of application and topic that was engaged at each moment (described later). The width of each vertical bar indicates the length of a unique smartphone session (screen on to screen off). Variability in temporal distribution of individuals' smartphone use is apparent when looking across persons, across days within person, and across hours within day.

Temporal aspects of individuals' smartphone use can be described at many different time-scales. Here, we illustrate two. At the daily time-scale, participants used their smartphones for, on average, 2.81 hours each day ($SD = 1.57$ hours, $Median = 2.64$ hours). Amount of use, however, also differed substantially across days. On the over 2,500 days of smartphone use observed in this study, total screen time ranged from a minimum of 5 seconds to a maximum of 13.61 hours, with the prototypical person's use varying substantially across days; average intraindividual standard deviation was 1.56 hours ($SD = 44.56$ minutes). All together the summaries highlight that total smartphone use varies substantially across days, both across persons and across days within person.

Zooming in, we can also quantify temporal features of use in terms of *session* – formally defined as the length of time that the smartphone was continually activated. Given our five-second sampling interval, sessions were demarcated by the screen having been turned off for ten or more seconds. Across 2,791 person-days of data collection, these 132 adults used their phones a total of 817,026 times (discrete sessions). The median number of sessions per day was 227.99 ($M = 275.00$, $SD = 201.14$), with the prototypical (median) session lasting 10 seconds ($M = 36.74$ seconds, $SD = 2.57$ minutes). Although the typical session was not very long, users often engaged many different applications in any given session (as seen in Figure 1 by the coloring of the vertical bars).

At the sample level, session length varied systematically within a day. The average session length during each hour in the day (midnight to midnight) is shown in Figure 2. Colored lines show how average session length changed across the day for a random subset of 30 participants to facilitate viewing of individuals' trajectories, and the thick black line shows how the sample-level average changed across the day for those with phone use during that hour. The curvature of the black line indicates an increase in average session length during the overnight hours, followed by relative stability in average session length across the daytime hours. We cautiously interpret this trend as indicating that session lengths increase overnight, but this interpretation is qualified by overall patterns in use. The background shading in Figure 2 displays how many screenshots were captured during each hour across all participants, with lighter colors indicating fewer screenshots captured during a particular hour. We see that the number of screenshots captured overnight decreases, so the increase in session length overnight may be driven by the small subset of participants who are actually on their smartphones very late at night/early in the morning. The across-person heterogeneity in session length and, especially, the hour-to-hour fluctuation in session length is prominently evident in the individual trajectories displayed in Figure 2. The extent of within-person variability in temporal aspects of smartphone use suggests the average trajectory is not actually representative of any individual's actual use pattern (ecological fallacy problem, see e.g., Molenaar, 2004). This phenomenon compounds as we delve into details of how and when individuals engage with other features of screen content.

4.2 Textual Content of Smartphone Use

Beyond the temporal aspects of smartphone use (which may also be available in some other logging paradigms), Screenomics provides an opportunity to examine the quantity of and speed at which content flows across individuals' smartphone screens.

Word count is a general measure of the amount of textual content that appears on individuals' smartphone screens. During a prototypical (median) session, across all participants and days, individuals encountered 64 words ($M = 271.47$, $SD = 1,000.60$), and during a prototypical day were exposed to 64,986.53 words ($M = 71,489.12$, $SD = 50,506.13$). There is high idiosyncrasy in the number of words an individual is exposed to in a given session (as seen by the large standard deviations). At the sample level, there was lower average word count for a given session during the overnight hours and relative stability across the daytime hours. This average is misleading, however, because it masks the actual heterogeneity of individuals' use. Furthermore, it is worth noting this quantification counts words that remain on the screen across multiple screenshots (e.g., while reading) multiple times, and thus does not consider how individuals are actually moving through the text that appears on their screen.

Word velocity is a measure of the speed at which the textual content on the screen changes (e.g., when scrolling through text). During a prototypical (median) session, across all participants and days, individuals encountered 8 ($M = 14.56$, $SD = 17.81$) new words every five seconds, and over the course of the entire day encountered 14.99 ($M = 15.46$, $SD = 5.53$) new words every five seconds.

Temporal distribution of average word velocity across the hours of the day is shown in Figure 3. Colored lines show how a random subset of 30 participants' average word velocity changed across the day, and the thick black line shows how the sample-level average changed across the day. The curvature of the black line indicates lower word velocity during the overnight hours, and relative stability across the daytime hours. Again, however, the sample-level trend must be qualified by the amount of use during each hour (background shading) and the extent to which word velocity differs across persons and across hours within each person. A preliminary interpretation is that this measure of textual content flow reflects changes in users' general activity levels (e.g., sleeping), specific activities (e.g., use of News & Magazines vs. Video Player applications), screen behaviors (e.g., scrolling speed), or cognitive load and engagement (e.g., arousal levels). Together, our simple quantifications of textual content and flow illustrate the value screenomes (screenshot sequences) have for the study of individuals' smartphone use and the cognitive processes and psychological motivations that drive that use.

4.3   Graphical Content of Smartphone Use

In addition to the textual content of individuals' smartphone screens, the Screenomics paradigm allows for examination of the graphical content of those screens.

Image complexity is a measure of how visually textured the graphical content of users' screens are, an indicator of visual stimulation. Generally, pictures and video are more textured and complex than text. During a prototypical (median) session, across all participants and days, image complexity was 3.40 ($M = 3.49$, $SD = 2.30$). Similarly, during a prototypical day, image complexity was 3.69 ($M = 3.63$, $SD = 0.85$). The relative homogeneity across time-scales and persons may be a function of the device and how content designers, across domains, take advantage of the smartphone display. Although the

graphical content of the screen is generally homogenous, how the graphical content flows across individuals' screens is a different story.

Image velocity is a measure of visual flow – i.e., how quickly images are moving on the smartphone screen. During a prototypical (median) session, across all participants and days, image velocity was 0.35 ($M = 1.01$, $SD = 1.57$), and during a prototypical day, image velocity was 1.24 ($M = 1.23$, $SD = 0.42$). These values indicate the graphical content of the screen is changing substantially every five seconds. The screens are by no means static. Temporal distribution of average image velocity across the hours of the day is shown in Figure 4. As with the other plots, cautious interpretation suggests that average image velocity changes systematically across the day (lower overnight). However, particularly prominent is the extent of within-person fluctuation in image velocity. Even at the hourly aggregation shown here it is clear that the flow of graphical content on individuals' smartphone screens is highly variable and fast-moving.

4.4 Topical Content of Smartphone Use

In addition to the general information of how textual and graphical content flow across individuals' smartphone screens, detailed information about the specific applications and topics users engage with over time can also be extracted from the screenome.

4.4.1   Applications—During the 8,000+ hours of in-situ smartphone use that we observed, these 132 users engaged with 315 unique applications. Of the 26 different types of applications represented in the data, the most commonly used ones were Social (33.63%, e.g., Facebook, Snapchat), Tools (25.32%, e.g., home-screen, notifications, other OS-type applications), and Communication (23.98%, e.g., texting) applications. The number of application types represented in each user's screenome ranged from 14 to 21 (*Median* = 17, $M = 17.58$, $SD = 1.54$). Temporal distribution of use was extremely heterogeneous, varying dramatically across persons and across time within each person. As evident in how the colors are distributed in Figure 1, there is very little systematicity in when and in what order individuals use all these applications. Multiple types of applications are often used in the same session, and there is a lot of switching among and between applications. Clearly, users are building and curating personalized threads of digital experience that cut very quickly between many different applications and platforms. They play games, listen to music, have social relationships, watch videos, and much more – all within seconds of each other.

4.4.2   Specific Topics—While the distribution of applications informs us about the types of activities that individuals engage on their smartphones over time, it does not provide information about the specific topics that individuals interacted with during those activities. Unable yet to obtain an exhaustive description of all the topics users engage, we illustrate how engagement with three specific topics, *nutrition, public affairs*, and *purchasing*, are encountered, no matter on what platform or application.

Generally, each of these three topics constitutes between 0.31% and 1.91% of individuals' smartphone use. However, this is actually somewhat substantial considering the hundreds of topics that individuals may encounter each month. Table 1 summarizes the occurrence of these three topics for the three participants depicted in Figure 1 and for the sample as a

whole. Temporal distribution of the three topics for the three participants is also shown in Figure 1, with the screenshots containing words from the nutrition, public affairs, and purchasing dictionaries indicated with black triangles (top quarter of each bar), squares (middle of each bar), and circles (bottom quarter of each bar), respectively. Importantly, Figure 1 indicates no direct correspondence between topics and type of application. For instance, content related to public affairs is encountered at all different times of day and in multiple types of applications.

In sum, the main take-home message from our analysis is that individuals' smartphone use is characterized by *individual, idiosyncratic movement.* Every screenome evidences the speed with which individuals are continually moving through and among a wide range of textual, graphical, and topical content.

5. Discussion

Much of the research on screen time and device use has used self-report, experience sampling, and application logging software to describe behaviors in the aggregate (Christensen et al., 2003; Delespaul, 1992; Reis & Gable, 2000; Cockburn & McKenzie, 2001; Jansen & Spink, 2006; Tossell et al., 2012; White & Huang, 2010). While these techniques have provided a base of knowledge about general use, there is growing recognition that research needs to move beyond *how* much time people spend on their devices, towards description and understanding of *what* individuals do on their devices (Reeves et al., 2020). In this study, we used Screenomics (Reeves et al., 2019; Ram et al., 2020) to describe temporal, textual, graphical, and topical features of what appears on people's screens.

We examined when and for how long individuals used their smartphone at daily and hourly time-scales. We found, at the aggregate level, that session length was fairly stable during the day, but increased late at night (longest at around 4 am). We speculate that this increase in session length is due to the types of media experiences that may occur late at night as opposed to at other times in the day (e.g., watching movies and scrolling through Instagram vs. responding to text messages). However, not all participants used their smartphone this late at night (as can be seen by the background shading in Figures 2 to 4), so this finding may be driven by a particular type (or subset) of smartphone user. Importantly, amount of use and fragmentation of use differed substantially both between persons and within persons over time. This aspect of individuals' smartphone use, which becomes more prominent when analyzing fine-grained data like ours, suggests that aggregate summaries of use have limited value when examining sample or individual use. In short, no one is actually using their smartphone in an "average" way, relative to others or relative to themselves. This has implications for designing device use tools that promote awareness of one's activities. As our data show, comparisons across individuals and the notion of "average time spent" provided in screen time summaries are probably misleading. Instead, within-person trends over time are likely to be more revealing and useful for meaningful interpretation of device use data. For the same reason, public policy recommendations regarding smartphone use that focus on comparisons to average use also lack relevance to the way that individuals are actually using their phones. These problems are further magnified when based on self-

reports (e.g., I spend > 3 hours per days on my phone, I watch a lot of videos on my phone every day).

We also examined how textual features (word count and velocity) and graphical features (image complexity and velocity) of phone content changed at the session, hourly, and daily time-scales. Prior work has examined the association of word count and image complexity with a variety of individual factors (e.g., social media engagement, Shami et al, 2015; depletion of cognitive capacity, Thorson et al., 1985). While our findings indicate aggregate level trends in word count (i.e., a dip in word count during the overnight hours) and aggregate level stability in image complexity throughout the day, our variability metrics illustrate these aggregate level averages mask the idiosyncrasy of individuals' exposure to word count and image complexity. From individuals' time-varying exposure to word count and image complexity, we might infer changing levels of engagement and cognitive capacity throughout the day (although future research needs to measure these outcomes directly). We also examined how exposure to words and images changed as individuals scrolled through content. Specifically, we found that, at the aggregate level, word and image velocity decreased late at night. Our emerging hypothesis is that individuals' cognitive load capacity decreases during late night hours as users become tired and are unable to process textual and graphical content as quickly. Alternatively, at least some subset of users interacts with more image-based materials during overnight hours (e.g., scrolling through Instagram vs. reading and writing text messages). Similar to applications that adjust screen brightness based on the time of day to accommodate sleep/wake cycles, our findings inform researchers and designers about how other aspects of media content can be titrated and matched to diurnal patterns in mobile engagement/information flow. For example, given that the complexity of visual information is associated with the depletion of cognitive capacity (Thorson et al., 1985), researchers might use our definition of image complexity to predict individuals' expenditure of cognitive capacity, thus informing the design of messages (text-heavy vs. visually stimulating) to be optimally effective.

Finally, we demonstrated how Screenomics can be used to examine specific types of content. As can be done with other logging paradigms, we described the different types of commercial applications that users engage. Most notable was how rapidly individuals switched from one application to another. The device clearly supports individuals' abilities to move quickly through lots of tasks and topics. Designers may consider further how to optimize users' desire to engage multiple types of applications in the same session as they actively curate their digital experience at high speed. We also examined all the text that our OCR module extracted from each screen to describe how users engaged with nutrition, public affairs, and purchasing-related topics. Contrary to the idea that individuals use specific apps for a specific purpose, we found that the topics (i.e., words in the three non-exhaustive dictionaries) were not platform-specific. Rather, those topics were present across a wide range of types of applications.

These results encourage further consideration of whether commercial categories represent how users actually use applications, and whether specific applications can be used in new ways to explore new topics. Usability and user-experience studies will likely be profoundly redefined by adopting a Screenomics approach by eschewing industry-based categorizations

of media content (such as type of application or platform, or visual or functional elements of interface design) in favor of more individualized, granular, psychologically relevant features of text, graphics, and time. This can help us discover those elements of interfaces and interactions that are primordial, such as specific words and colors, especially in combination with time of day, content, and surrounding contexts, and motivate design efforts around such elements.

5.1   Privacy Considerations

The collection of participants' screenomes provides a rich repository of digital life and new insights about how individuals engage with digital media in everyday life. These collections and data also require substantial consideration of a variety of privacy issues and protocols that support and maintain confidentiality. This project passed a review by the Institutional Review Board overseeing human subjects research at our institution in coordination with the Office of Privacy and Office of Information Security. Those reviews included university faculty, administrators with extensive experience in research with human subjects, system engineers, and legal consultants that considered all the relevant consent, privacy protection, and legal and ethical issues. We continue to monitor and consider best practices of data security and privacy relevant to our project. Screenome data often contain highly personal or sensitive information, and we have designed our data collection and analysis protocol to maintain participants' privacy to the best of our abilities. During data collection, screenshots are encrypted and sent to a secure server only accessible by trained research staff on selected computers that are permanently located in a limited-access laboratory. During data analysis, we attempt to minimize the number of screenshots viewed by members of the study team by utilizing computational tools that can automatically extract features from screenshots (e.g., OCR) and machine learning methods that rely on only small batches of human-labeled data to propagate larger batches of machine-labeled (unseen) data. We also minimize interest in participants' personal or sensitive information through analysis of selected features that obscure those details, for example by only extracting summary textual and graphical information (e.g., word count, image velocity) rather than more specific details. Future studies using the Screenomics framework should work to develop a data collection and analysis pipeline that simultaneously takes advantage of the richness of Screenomics data while preserving participant privacy.

5.2   Limitations and Future Directions

While the Screenomics framework provides many new opportunities to study when people use their smartphones and what they use their smartphone for, the approach and our initial data analyses have some limitations. First, our sample of users was a convenience sample that is younger and more urban than the general population of the United States. Our sample's smartphone use may not reflect smartphone use of the general population in amount of time, application, and content of use. Future work should apply Screenomics to a more representative sample, including children and adolescents, older adults, and residents of rural areas. However, even then, some care must be taken with regard to interpretation of aggregates and averages.

Second, our temporal sampling of individuals' smartphone screen was fast, every five seconds the screen was activated, but may not be fast enough. Some behaviors, such as scrolling, dismissing a notification, or checking the time, may have been missed because they take less than five seconds. Future collection of screenomes should consider sampling at even faster rates in order to capture another layer of behavior.

Third, the OCR text extraction that serves as the basis for computation of textual features and dictionaries has imperfect precision (accuracy = 74%), meaning that we may be missing some words and/or getting a few extras. Furthermore, the dictionaries we used to operationalize specific topics (e.g., public affairs) were non-exhaustive, and thus might have missed relevant screenshots that contained related words or images. Future research should work to both understand how these measurement limitations influence results and improve on the algorithms used for recognizing characters and words that appear on heterogenous and complex backgrounds. There is much room to expand the set of textual and graphical features to include, for example, recognition of faces, logos, icons, objects, and the content/meaning of images (e.g., Kumar et al., 2020).

Fourth, we recognize that reactivity effects – i.e., participants changing their behavior in response to being observed – may be a consequence of the Screenomics approach. Prior research (e.g., Yeykelis, Cummings, & Reeves, 2018) observed users' laptop use for four days and found that users continued to participate in behaviors that may be considered less socially desirable (e.g., watching pornography). We acknowledge that we do not know the full extent of reactivity effects and further work is needed to examine if and how participants change their behaviors knowing that the study software is capturing screenshots of their device (e.g., by observing and comparing use patterns at the beginning versus the end of the study time period when the effects of reactivity may be different or by directly asking participants after study completion if they were aware of the software when using their devices or changed their behavior).

Fifth, we only reported on participants' smartphone use in this study; however, people often use and switch between multiple devices every day (Reeves et al., 2019). Future work should expand the number of screens observed to include laptops, tablets, and televisions.

Sixth, the benefits of the current Screenomics system and analysis pipeline to the study participants were limited to monetary compensation for participation. The foundations conveyed here, though, are intended to support collection and interactive delivery of screen information that will inform individuals about their technology use. Eventually opt-in interventions that use information embedded in the screenome will help individuals optimize their personal goals, health, and well-being. For example, our team is currently developing interactive dashboards that allow users to obtain greater insight into their screen use behaviors.

Finally, using a Screenomics framework requires substantial time and computing power to process and analyze the screenshot data. We acknowledge that URL and application logging tools may be simpler and sufficient alternatives to Screenomics for some uses/research questions; however, application logging tools also miss or purposively disregard many

aspects of the specific content and other details appearing on people's screens. That being said, we are working towards integration of these tools with the newly available Screenomics streams. For example, a logging tool can provide an OS-derived label of the current front-facing application, greatly reducing (but not entirely eliminating) the need for human coders. We believe the information gained using Screenomics is highly valuable and that our team has already begun laying the groundwork for this approach for other researchers to build upon. As this body of work expands, it will be important to continually update the screenome with new features that capture and describe relevant aspects of an ever-changing digital landscape.

6. Conclusion

In summary, this study has demonstrated how Screenomics can be used to describe what people see and do on their smartphones. We explored how textual and graphical information on individuals' smartphone screens change at multiple time-scales and highlighted how there is substantial heterogeneity of smartphone use and content both across persons and within persons across time, when observed and studied in fine-grained detail. We demonstrate the utility of Screenomics for describing and studying how individuals construct their digital lives, and have laid the foundation on which to build Screenomics at scale. With more accurate and detailed descriptions of everything individuals see and do on their screens, researchers can test and potentially improve upon current theories that address the underlying mechanisms and effects of media use, and develop new theories premised on seeing digital media at the resolution of the screenome. In turn, improved understanding of media use can lead to better predictions of media use behaviors, resulting in novel and precise interventions (e.g., mHealth) to help individuals benefit from their screen use and limit adverse impacts, and produce personalized content (e.g., changing the flow of content based upon individuals' use patterns). Insights from this screen-based, individual data-driven approach can help advance behavioral research by revealing systematic patterns in individuals' daily digital lives and by helping us revise and update theories of technology use that were based on limited knowledge of aggregate summaries. New details are within reach.

Acknowledgments

References

Andrews S, Ellis DA, Shaw H, & Piwek L (2015). Beyond self-report: Tools to compare estimated and real-world smartphone use. PLOS ONE, 10(10), e0139004. doi:10.1371/journal.pone.0139004 [PubMed: 26509895]

Böhmer M, Hecht B, Schöning J, Krüger A, & Bauer G (2011). Falling asleep with Angry Birds, Facebook and Kindle. Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI '11. doi:10.1145/2037373.2037383

Brubaker JR, Kivran-Swaine F, Taber L, & Hayes GR (2012). Grief-stricken in a crowd: The language of bereavement and distress in social media. In Sixth International AAAI Conference on Weblogs and Social Media.

Chiatti A, Yang X, Brinberg M, Cho MJ, Gagneja A, Ram N, Reeves B, & Giles CL (2017). Text extraction from smartphone screenshots to archive in situ Media Behavior. In Proceedings of the Knowledge Capture Conference - K-CAP 2017. doi:10.1145/3148011.3154468

Christensen TC, Barrett LF, Bliss-Moreau E, Lebo K, & Christensen TC (2003). A Practical Guide to Experience-Sampling Procedures. Journal of Happiness Studies, 4(1), 53–78. doi:10.1023/a:1023609306024

Cockburn A, & McKenzie B (2001). What do web users do? An empirical analysis of web use. International Journal of Human-Computer Studies, 54(6), 903–922. doi:10.1006/ijhc.2001.0459

Delespaul PAEG (1992). Technical note: devices and time-sampling procedures. The Experience of Psychopathology, 363–374. doi:10.1017/cbo9780511663246.033

Gilmore RO (2016). From big data to deep insight in developmental science. Wiley Interdisciplinary Reviews: Cognitive Science, 7(2), 112–126. doi:10.1002/wcs.1379 [PubMed: 26805777]

Hancock J, Liu X, French M, Luo M & Mieczkowski H (2019). Social media use and psychological well-being: a meta-analysis. Paper presented at 69th Annual Conference of International Communication Association, Washington DC.

Hastie T, Tibshirani R, & Friedman J (2009). The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media.

Jansen BJ, & Spink A (2006). How are we searching the World Wide Web? A comparison of nine search engine transaction logs. Information Processing & Management, 42(1), 248–263. doi:10.1016/j.ipm.2004.10.007

Judd CM, Westfall J, & Kenny DA (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. Journal of Personality and Social Psychology, 103(1), 54–69. doi:10.1037/a0028347 [PubMed: 22612667]

Kaufman DR, Patel VL, Hilliman C, Morin PC, Pevzner J, Weinstock RS, … Starren J (2003). Usability in the real world: assessing medical information technologies in patients' homes. Journal of Biomedical Informatics, 36(1–2), 45–60. doi:10.1016/s1532-0464(03)00056-x [PubMed: 14552846]

Kumar S, Ramena G, Goyal M, Mohanty D, Agarwal A, Changmai B, & Moharana S (2020). On-device information extraction from screenshots in form of tags. Proceedings of the 7th ACM IKDD CoDS and 25th COMAD. doi:10.1145/3371158.3371200

Liaw A, & Wiener M (2002). Classification and regression by randomForest. R news, 2(3), 18–22.

Molenaar PCM (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. Measurement: Interdisciplinary Research & Perspective, 2(4), 201–218. doi:10.1207/s15366359mea0204_1

Murnane EL, Abdullah S, Matthews M, Kay M, Kientz JA, Choudhury T, … Cosley D (2016). Mobile manifestations of alertness. Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '16. doi:10.1145/2935334.2935383

Ram N, Yang X, Cho M-J, Brinberg M, Muirhead F, Reeves B, & Robinson TN (2020). Screenomics: A new approach for observing and studying individuals' digital lives. Journal of Adolescent Research, 35(1), 16–50. doi:10.1177/0743558419883362 [PubMed: 32161431]

Reeves B, Ram N, Robinson TN, Cummings JJ, Giles CL, Pan J, … Yeykelis L (2019). Screenomics: A framework to capture and analyze personal life experiences and the ways that technology shapes them. Human–Computer Interaction, 1–52. doi:10.1080/07370024.2019.1578652

Reeves B, Robinson T, & Ram N (2020). Time for the Human Screenome Project. Nature, 577(7790), 314–317. doi:10.1038/d41586-020-00032-5 [PubMed: 31942062]

Reeves B, Thorson E, & Schleuder J (1986). Attention to television: Psychological theories and chronometric measures. Perspectives on media effects, 251–279.

Reeves B, Yeykelis L, & Cummings JJ (2015). The use of media in media psychology. Media Psychology, 19(1), 49–71. doi:10.1080/15213269.2015.1030083

Reis HT, & Gable SL (2000). Event-sampling and other methods for studying everyday experience Handbook of research methods in social and personality psychology, 190–222. Cambridge University Press: New York, NY.

Roehrick K, Brinberg M, Ram N, & Reeves B (2019). The screenome: An exploration of smartphone psychology across place and person. Paper presented at the annual Society for Personality and Social Psychology conference, Portland, OR.

Shami NS, Muller M, Pal A, Masli M, & Geyer W (2015). Inferring employee engagement from social media. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15. doi:10.1145/2702123.2702445

Shannon CE (1948). A mathematical theory of communication. Bell system technical journal, 27(3), 379–423.

Sherry JL (2004). Flow and media enjoyment. Communication theory, 14(4), 328–347.

Sundar SS, Bellur S, Oh J, Xu Q, & Jia H (2013). User experience of on-screen interaction techniques: An experimental investigation of clicking, sliding, zooming, hovering, dragging, and flipping. Human–Computer Interaction, 29(2), 109–152. doi:10.1080/07370024.2013.789347

Sundar SS, & Limperos A (2013). Uses and grats 2.0: New gratifications for new media. Journal of Broadcasting & Electronic Media, 57 (4), 504–525. doi:10.1080/08838151.2013.845827

Tausczik YR, & Pennebaker JW (2009). The psychological meaning of words: LIWC and computerized text analysis methods. Journal of Language and Social Psychology, 29(1), 24–54. doi:10.1177/0261927x09351676

Thorson E, Reeves B, & Schleuder J (1985). Message complexity and attention to television. Communication Research, 12(4), 427–454. doi:10.1177/009365085012004001

Tossell C, Kortum P, Rahmati A, Shepard C, & Zhong L (2012). Characterizing web use on smartphones. Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems - CHI '12. doi:10.1145/2207676.2208676

van Berkel N, Luo C, Anagnostopoulos T, Ferreira D, Goncalves J, Hoslo S, & Kostakos V (2016, 5). A systematic assessment of smartphone usage gaps. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (pp. 4711–4721).

Visch V, & Tan E (2007). Effect of film velocity on genre recognition. Media Psychology, 9(1), 59–75. doi:10.1080/15213260709336803

Wang Z, Lang A, & Busemeyer JR (2011). Motivational processing and choice behavior during television viewing: An integrative dynamic approach. Journal of Communication, 61(1), 71–93. doi:10.1111/j.1460-2466.2010.01527.x

White RW, & Huang J (2010). Assessing the scenic route: measuring the value of search trails in web logs. In Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval (pp. 587–594). ACM.

Yeykelis L, Cummings JJ, & Reeves B (2014). Multitasking on a single device: Arousal and the frequency, anticipation, and prediction of switching between media content on a computer. Journal of Communication, 64(1), 167–192. doi:10.1111/jcom.12070

Yeykelis L, Cummings JJ, & Reeves B (2018). The fragmentation of work, entertainment, e-mail, and news on a personal computer: Motivational predictors of switching between media content. Media Psychology, 21(3), 377–402. doi:10.1080/15213269.2017.1406805

- Descriptions of media use based on aggregate individuals' experiences may be wrong

- Screenomics provides in situ observation of individuals' actual digital life

- Textual, graphical, topical content of individuals' screens differs substantially

- The details in the screenome open new insights for researchers/designers

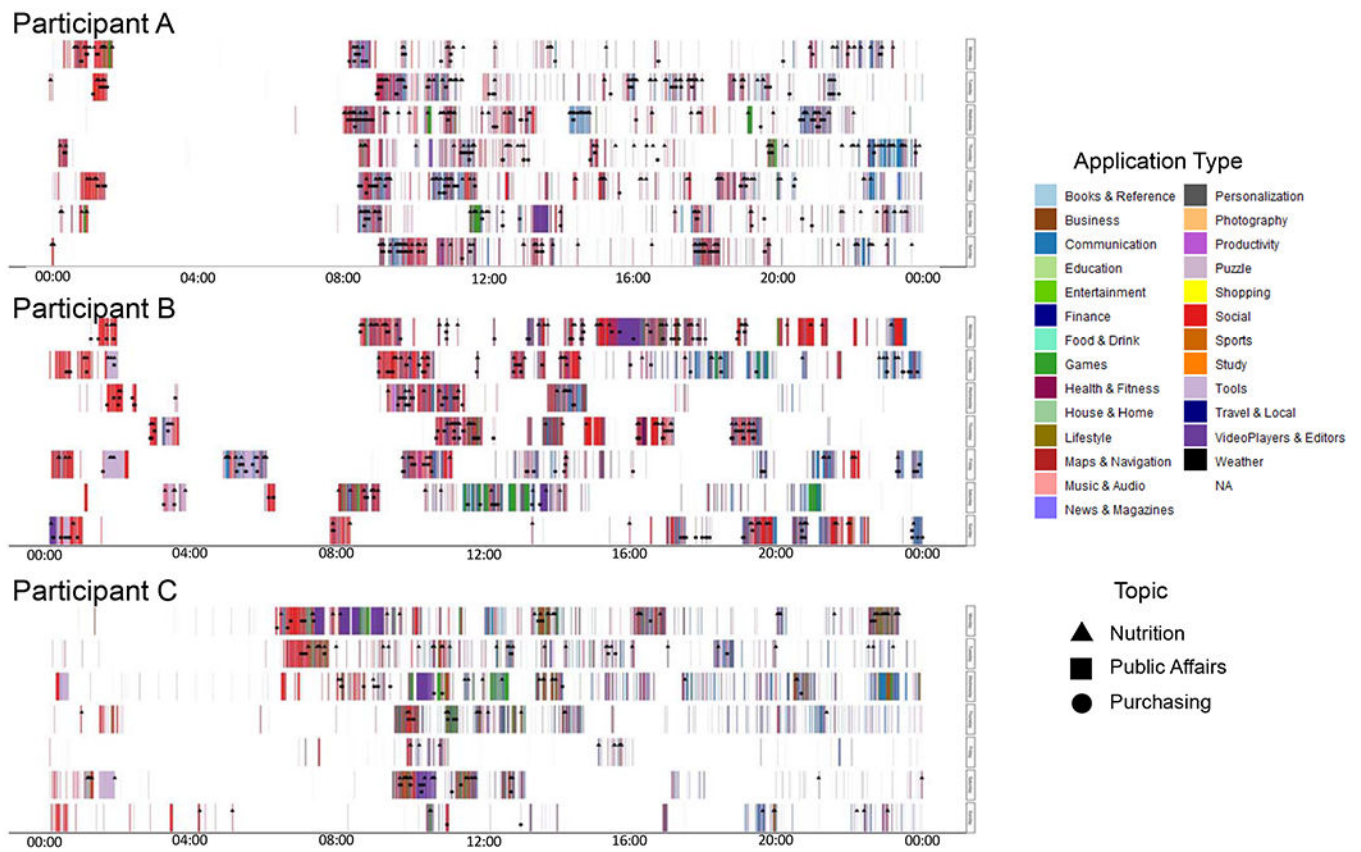- Observation at fine resolution suggests need for new theories of media use

Figure 1.

Three participants' screenomes – where the rows indicate smartphone use during each of 7 consecutive days, from Midnight (left) to 23:59 PM (right) on Monday (top) to Sunday (bottom). Colored vertical bars indicate when the smartphone was in use and which application category was on the screen. Shapes and the locations of the shapes on the colored vertical bars indicate when words from the nutrition, public affairs, and/or purchasing dictionaries were present on the screen.
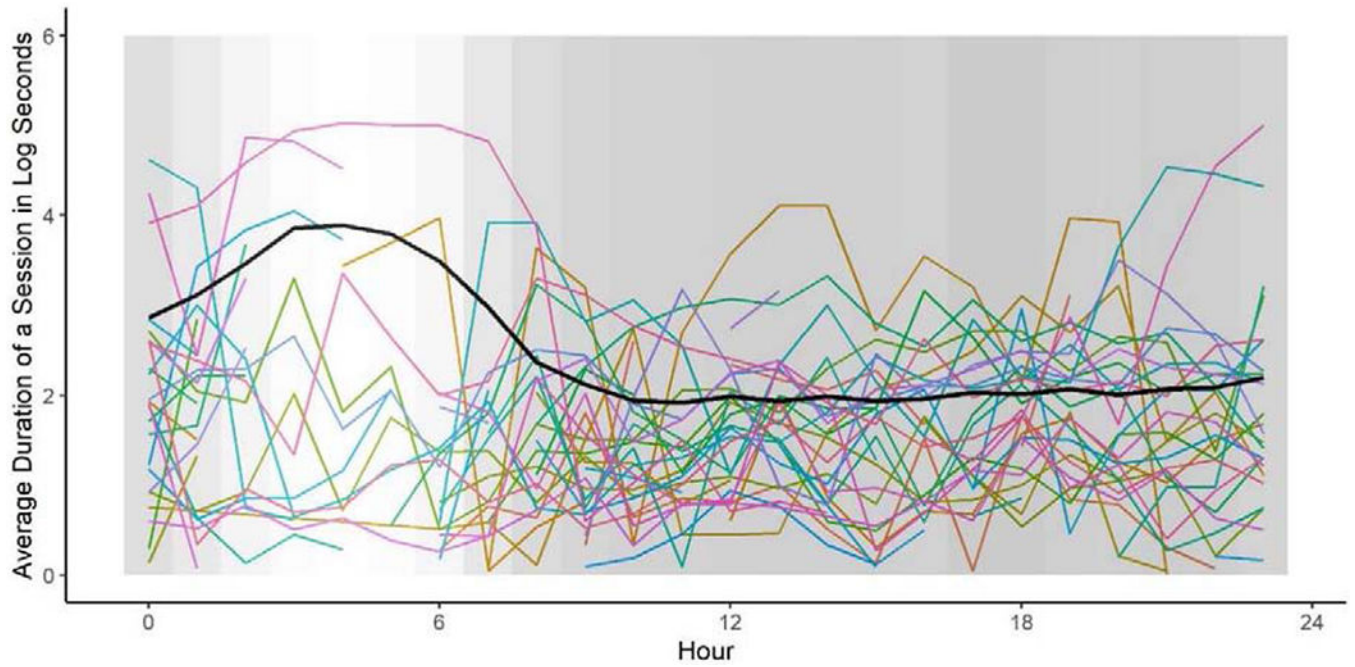
Figure 2.
Average session length (in log seconds) hour-by-hour, from Midnight (left) to 23:59 PM (right). The black line represents the sample level average for those with phone use during those hours, and each colored line represents a participant's average (subset of $n = 30$ participants to facilitate viewing of individuals' trajectories). The background shading displays how many screenshots occurred during each hour across all participants, with the lightest color indicating the minimum number of screenshots ($n = 74,297$) and the darkest color indicating the maximum number of screenshots ($n = 333,703$).
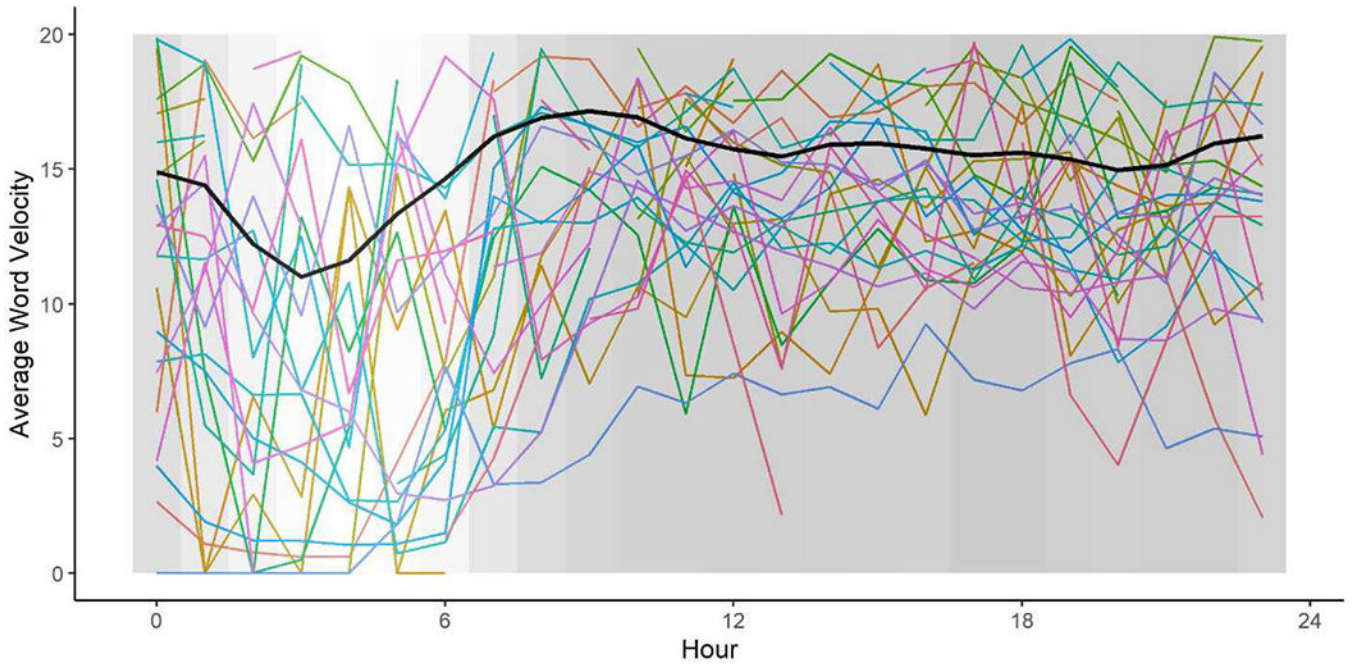
Figure 3.
Average word velocity (in log seconds) hour-by-hour, from Midnight (left) to 23:59 PM
(right). The black line represents the sample level average for those with phone use during
those hours, and each colored line represents a participant's average (subset of $n = 30$
participants to facilitate viewing of individuals' trajectories). The background shading
displays how many screenshots occurred during each hour across all participants, with the
lightest color indicating the minimum number of screenshots ($n = 74,297$) and the darkest
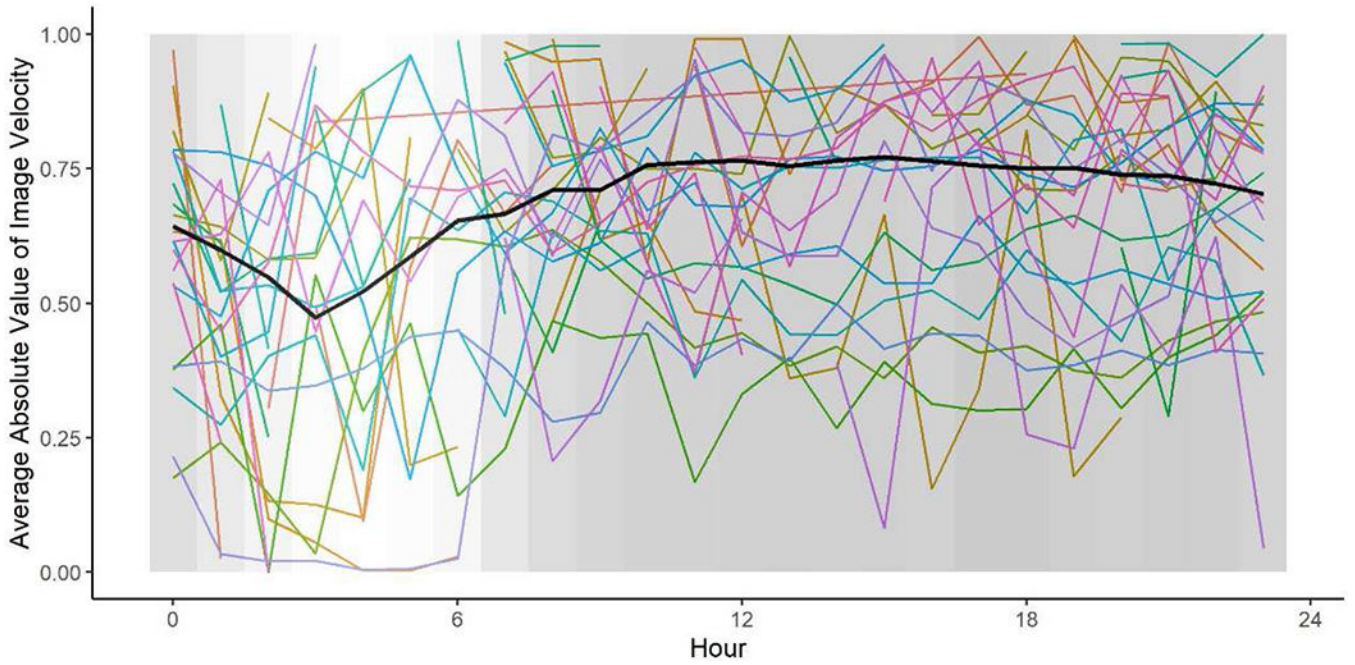color indicating the maximum number of screenshots ($n = 333,703$).

Figure 4.

Average absolute value of image velocity (in log seconds) hour-by-hour, from Midnight (left) to 23:59 PM (right). The black line represents the sample level average for those with phone use during those hours, and each colored line represents a participant's average (subset of $n = 30$ participants to facilitate viewing of individuals' trajectories). The background shading displays how many screenshots occurred during each hour across all participants, with the lightest color indicating the minimum number of screenshots ($n = 74,297$) and the darkest color indicating the maximum number of screenshots ($n = 333,703$).

Table 1

Percentage of topic occurrence for each participant and entire sample.

	% Nutrition	% Public Affairs	% Purchasing
Participant A	8.44	3.74	0.72
Participant B	2.56	1.35	1.48
Participant C	1.55	0.39	0.06
Total Sample Means	1.91	0.76	0.31
Total Sample Standard Dev.	1.46	0.94	0.35
Range	0.03 – 8.17	0.00 – 5.21	0.00 – 1.78