Article

mmc

Mobile Media & Communication
2022, Vol. 10(3) 487–509

DOI: 10.1177/20501579221080333
journals.sagepub.com/home/mmc
SAGE

Selectively localized: Temporal and visual structure of smartphone screen activity across media environments

Daniel Muise, Yingdan Lu, Jennifer Pan, and Byron Reeves
Department of Communication, Stanford University, USA

Abstract
This study demonstrates how localization and homogenization can co-occur in different aspects of smartphone usage. Smartphones afford individualization of media behavior: users can begin, end, or switch between countless tasks anytime, but this individualization is shaped by shared environments such that smartphone usage may be similar among those who share such environments but contain differences, or *localization*, across environments or regions. Yet for all users, smartphone screen interactions are bounded and guided by nearly identical smartphone interfaces, suggesting that smartphone usage may be similar or *homogenized* across all individuals regardless of environment. We study homogenization and localization by comparing the temporal, visual, and experiential composition of screen activity among individuals in three dissimilar media environments—the United States, China, and Myanmar—using one week of screenshot data captured passively every 5 s by the novel Screenomics framework. We find that overall usage levels are consistently dissimilar across media environments, while metrics that depend more on moment-level decisions and user-interface design do not vary significantly across media environments. These results suggest that quantitative research on homogenization and localization should analyze behavior driven by user interfaces and by contextually determined parameters, respectively.

Keywords
smartphone, screenomics, screen, localization, homogenization, session, entropy

Corresponding author:
Yingdan Lu, 450 Jane Stanford Way, Bldg 120, Rm 110, Stanford, CA, 94305, USA.
Email: yingdan@stanford.edu.

Introduction

At any moment, smartphone users can begin, end, or switch between countless tasks on a single screen (Goggin, 2010). However, individuals with shared environments are more likely to share in the characteristics of their media behavior; hence, individualization implies *localization* of smartphone usage style, with individuals sharing more in common with group members than with out-group members (Keller, 2002; Van Alstyne & Brynjolfsson, 2005). However, all smartphone users' screen interactions are bounded and guided by nearly identical smartphone interfaces (Sundar & Oh, 2019). This may constrain device use, making smartphone usage style *homogenized* across individuals regardless of shared traits or environments (Jenkins, 2004; Punathambekar & Mohan, 2019). Hence, our devices present a paradoxical duality: they may amplify both uniqueness *and* similarity across groups.

We shed light on this duality at a granular level by comparing the temporal and visual composition of screen activity among smartphone users in three dissimilar media environments—the United States, China, and Myanmar—using one week of screenshot log data captured passively every 5 s that smartphone devices are turned on. Specifically, we establish for the first time that homogenization and localization act simultaneously on different factors of smartphone use. To do so, we make use of a basic insight: all smartphone behaviors logically lie on a spectrum ranging from the most to least associated with local influences and constraints, such as media infrastructure, device availability, economic wealth, and political censorship. Hereafter, we refer to these influences and constraints collectively as *contextually determined parameters*. We do not aim to enumerate these parameters nor make claims about any specific countries, but we do assume that strong variations in these parameters taken together are more likely to moderate some aspects of screen activity (e.g., total screentime in a week) than aspects that are based on moment-level decisions and even user-interface design (e.g., the average duration of a usage session). Our analysis finds that the former set of metrics indeed shows pronounced dissimilarity across environments (i.e., localization), while the latter does not.

Our analysis expands the literature with several methodological innovations. Our screenshot-based log data enable an analysis of media usage that cuts across applications and tasks (Reeves et al., 2021). We develop this asset into metrics that summarize the temporal and visual features of smartphone usage, including the as-yet uninvestigated relationship between visual entropy and temporal "depth" in a usage session. Finally, to automatically sort millions of smartphone screenshots into meaningful visual categories, we synthesize image-based unsupervised machine learning, deep learning, and human validation.

The rest of the paper is structured as follows. The second section offers a theoretical framework for understanding localization and homogenization. The third section describes our methodological approach and the details of our data and sample. The fourth section describes our results, and the fifth section discusses and contextualizes our findings.

Research background & theoretical framework

Localization

Broadly, localization refers to how media technology's affordances are adapted in different ways by different groups of people (Mansell, 2012, 2017; Postill, 2008; Price, 2002; Schiller, 1976). In its most basic form, this can be measured in overall usage. People use technology more or less often according to how *useful* the technology is to them personally (Bagozzi, 2007; Bagozzi et al., 1992), and that usefulness is moderated by a host of contextually determined parameters, such as the amount of accessible content, the utility of the device to a user's particular needs, and social customs and manners (e.g., Campbell, 2007; Mante, 2002). In other words, boundaries in society influence how we make use of device flexibility (Kappos & Rivard, 2008).

The more that users are able to individualize how they use a technology, the more that traits specific to a group can manifest in usage behavior, assuming commonalities among members of a group and dissimilarities between groups (Keller, 2002; Van Alstyne & Brynjolfsson, 2005). The smartphone in particular is a prime candidate technology for localization, due to its unprecedented flexibility and the implications of its flexibility for individualization of usage styles (Kang & Sundar, 2016; Reeves et al., 2021; Turow, 2011). As with any digital computer, smartphone devices are meta-media; they afford the representation and recombination of other forms of media (Kay & Goldberg, 1977). The Internet, smartphones, and social media have layered experience into increasingly complex and idiosyncratic combinations of form and content (Jensen, 2011). Mediated tasks have become more fragmented as users actively weave together radically different short segments. For example, smartphone users can stop a movie to answer a text for 10 s, then play a game before resuming the movie (Reeves et al., 2021).

Under the localization paradigm, researchers in mobile communication have compared self-reported smartphone usage across distinct groups of people, finding that location is connected to content selection, self-reported screentime, and mobile phone etiquette (Baron & Segerstad, 2010; Campbell, 2007; Mante, 2002). However, localization can be deceptively complex to measure granularly, owing again to the flexibility of the smartphone. Consider these examples from our corpus of screenshot logs (detailed in full in the third section), each capturing roughly an hour of usage on a Saturday night in 2017 in very different locations:

> On May 27th at 9:32 PM, subject #20 in Los Angeles turned on the phone screen, switched away from a YouTube hip-hop playlist, and opened up Instagram. An auto-repair company's ad on Instagram led to three minutes of examining the company's profile, before the user searched for a residential address on Google, spending three minutes evaluating the real estate of the address via mobile browser.

> On July 15th at 9:53 PM, subject #12 in Yangon, Myanmar, played the game "Clash of Clans" for one minute until the game application crashed, then opened Facebook Messenger and checked new messages, replied to a goodnight text, played "Clash of Clans" for 12 minutes, opened Facebook, checked general notifications, then navigated to

> the Facebook News Feed, and spent several more minutes browsing photos and other apps before shutting down for the night.

> On August 19th at 9:00 PM, subject #51 in China spent 10 minutes in a group chat on WeChat before navigating to Sogou Browser and then began to read an e-book. Four minutes later, a notification from a friend via WeChat about an ongoing sale initiated 10 minutes of online shopping on the Taobao shopping app, followed by 10 minutes of discussion with customer service, before spending hours reading the e-book.

These vignettes demonstrate how a measurement of localization can depend undesirably on the level of analysis implicitly selected by the researcher. For example, the subjects each used different particular apps that are common to their local context, but the categories of apps they used were fairly universal. Hence, we approach our localization measures very carefully by focusing mainly on the temporal and visual structure of usage. *Temporal* structural features of smartphone usage have been used to study screen addiction ((Lin et al., 2015), attention and multitasking (Oulasvirta et al., 2005), and arousal in relation to the screen (Yeykelis et al., 2014). Here they include overall time spent using the phone, average duration of screen inactivity, and total number of usage sessions. Each of these three features reflects overall usage, which we expect to differ across locations in line with contextually determined parameters that influence the utility of the smartphone. *Visual* structure refers to the arrangement of pixel hue and brightness on the screen; we measure it here by the *image entropy* (Shannon, 1948; Yang et al., 2019; Zheng et al., 2009).[1] We expect image entropy to vary across locations when device usage is driven by the user rather than the interface. We proxy that difference here by examining usage occurring after many seconds of screen activity, thus excluding quick interface-driven behavior like checking the screen for notifications.

These four structural features are commensurable across locations and agnostic to particular applications and levels of analysis, and lead us to our first set of hypotheses:

H1. Smartphone usage behaviors are expected to differ across geographically distinct groups of smartphone users in both temporal and visual structural features, including:

H1a. Overall screen time;

H1b. Total number of smartphone sessions (equivalently, "screen unlocks");

H1c. Average time between smartphone usage sessions;

H1d. Image entropy of screenshots captured later in continuous smartphone usage.

Localization also applies to user engagement with content and activities. The vignettes above demonstrate that particular applications or behaviors are a poor way to measure and compare user activities, but the prevalence of broad categories of activity can be usefully compared across groups. One such category that we anticipate to be localized is social media. While all subjects in our study used social media in some form (reflecting a trend among virtually all smartphone users), the manner in which social media fit into their lives depended on the affordances of the apps preferred in their group. This leads to an expectation of localization in the degree to which social media is used. Therefore, we have our next hypothesis:

H2. Smartphone usage behaviors are expected to differ among groups of smartphone users sampled from Myanmar, China, and the United States in their use of social media.

Homogenization

Although the selection and sequencing of media content are undoubtedly more individualized than ever, the manner in which users access content is increasingly homogeneous (Humphreys et al., 2018). Smartphones are built according to standardized designs (e.g., standard iOS systems and hardware parameters and design of iPhones) that are intended to elicit standardized behaviors in consumers (Jenkins, 2004; Mansell, 2017). Hence, homogenization of smartphone usage style is driven by the top-down standardization of smartphones intended to streamline and optimize user experience (Goggin, 2010), and by bottom-up user adaptation to the smartphone's affordances for the universal goal of easy communication (Smith & Kantor, 2008). Companies are economically incentivized to continually simplify the smartphone format into forms that are intuitive globally, and these standardized designs scale as companies acquire larger shares of the global market (Goggin, 2010; Langley & Leyshon, 2017; Price, 2002; Sparks, 2012; West, 2019). Standardization of media format shrinks the scope of how information is produced and consumed (Goggin, 2010). Thus, perhaps counterintuitively, the advent of flexible devices that superficially enable fragmentation of tasks and content may catalyze *homogenization* (Mansell, 2012, 2017).

In behaviors that are more closely tied to the standardized interface, we do not expect to find differences across distinct groups. That is, there are universal characteristics of smartphone usage that we do not expect to be moderated by contextually determined parameters, but instead to be relatively homogeneous. As with our examination of localization, we focus on structural features of smartphone usage. Here we focus on the average duration of smartphone usage sessions and the relative frequency of usage sessions of varying durations. Note that smartphone sessions are fleeting time spans measurable in seconds, often occurring out of habit, and rarely intended to last for a predetermined period of time (Reeves et al., 2021). Hence, the average session duration is not dependent on overall usage time nor overall utility, and so we do not expect such a feature to be moderated by location (Shepard et al., 2011). Rather, such a behavior would reflect homogenization. We also expect the visual structure of smartphone usage to be homogeneous across distinct groups during usage that is likely driven by the interface. Mirroring our localization hypothesis H1d, we proxy interface-driven behavior by comparing screen activity occurring early in smartphone usage sessions, and proxy visual structure with image entropy. Therefore, we have our next hypothesis:

H3. Smartphone usage behaviors are not expected to have significant differences among groups of smartphone users sampled from Myanmar, China, and the United States in:

H3a. Average session duration;
H3b. Rapidity of session durations, measured by the power-law exponent that best matches their intra-individual distribution;
H3c. Image entropy of screenshots captured early in continuous smartphone usage.

Mirroring our expectation of localization of social media prevalence in H2, we expect relative homogeneity in activities more closely related to the interface and elicited behaviors. Here we focus on typing. Typing behavior reflects the level of interactivity and production behaviors during smartphone sessions (Cho, 2020). The interface tools by which smartphone users produce text (i.e., the keyboard) are strikingly uniform across devices, software, and alphabets, and we do not expect structured differences in the desire or ability to produce textual content or to communicate interpersonally.

H4: Smartphone usage behaviors are not expected to have significant differences among groups of smartphone users sampled from Myanmar, China, and the United States in typing behaviors.

Data and method

Data

Our data consist of week-long screenshot corpi from smartphone users in Myanmar, China, and the United States, three dramatically different smartphone environments in terms of economic resources, technological history, language, and more (Ling et al., 2017; Mansell, 2017; Price, 2002; Yu, 2017). These distinct contexts are suited to detect localization in smartphone usage and provide a hard test in a search for homogeneity.

Myanmar's Internet and telephone penetration rates rapidly increased from near-zero levels to .9 mobile subscriptions per capita from 2012 to 2017 (International Telecommunication Union [ITU], n.d.). However, local content was relatively scarce and suffered from non-standardized language encoding (Leong, 2017). In 2017, China had a 48% smartphone penetration rate (Statista, 2019a), a flourishing domestic app ecosystem, an extensive online censorship regime (King et al., 2013, 2014; Wu & Taneja, 2016), and a broad array of uses for the smartphone (Taneja & Wu, 2014). The leading Chinese social networking app, WeChat, dominates as a mobile services hub and communication platform (Montag et al., 2018). The United States, with smartphone penetration at 76% (Statista, 2019b), has long led media design (Hoskins & Mirus, 1988), and its technology and content have diffused globally (Price, 2002; Schiller, 1976). Smartphone infrastructure in China and Myanmar reflects this diffusion: from the design of the smartphone to the standardized nature of social media, and in Myanmar, a U.S.-based app ecosystem centered on Facebook (Leong, 2017). Meanwhile, China has grown its cultural and technological influence over neighbors (including Myanmar), with Chinese smartphones flooding the market (Leong, 2017; Sparks, 2012). The result is that these three states host mutually distinct smartphone environments, shaped by unique combinations of political and economic forces.

Subjects in the United States were recruited from greater Chicago, New York, and Los Angeles. Subjects in Myanmar were recruited from greater Yangon. Subjects from China were university students from a number of metropolitan areas. Demographics and summary statistics are displayed in Table 1. The statistical tests that follow leverage the variation associated with the media environments of these three sampling sources in order to study homogenization and localization in smartphone usage. We emphasize that our intent is *not* to treat these samples, or our conclusions, as representative of any of these three nations, and our findings do not describe the experience of the typical smartphone user in any country. The specific locations or populations from which we sample our subjects are not of direct relevance to our outcomes.

Our screenshot data were collected using the Screenomics framework (Reeves et al., 2021). As approved by our Institutional Review Board, subjects installed a custom-built application on their own smartphone devices.[2] The software passively captures, encrypts, and uploads a time-stamped screenshot every 5 s as subjects go about their daily lives. Though few studies have used logs to examine cross-national variation in smartphone usage (Mehrotra et al., 2017; Qin et al., 2018), smartphone logs allow unprecedented granularity and applicability when studying subtle behaviors (Harari et al., 2016), far beyond what self-reports can capture (Boase & Ling, 2013; Kobayashi & Boase, 2012; Reeves et al., 2020), without requiring the artificiality of laboratory research (Reeves et al., 1986; Rothschild et al., 1986). Screenshot collections go beyond other logging techniques because screenshots allow content analysis and visual analysis that is agnostic to app choices (Brinberg et al., 2020).

Demographic information was collected via surveys: all subjects are over 18 years of age and each is the sole user of only one Android smartphone. Each subject installed the application in mid-2017. One continuous week of data is selected from each individual that began at least 30 min after app installation and ended 7 days thereafter. Figure 1 visualizes one day of activity from each subject, illustrating the granularity and variability of smartphone usage. Due to the complexity of recruiting subjects into this research, our cross-sectional N is relatively small; we address this limitation statistically with multiple robustness checks and discuss limitations in the Conclusion section.

Table 1. Sample summary statistics at the subject level and screenshot level

	China	**Myanmar**	**United States**
n	47	19	82
Age (minimum, median, maximum)	19, 21, 34	20, 28, 35	21, 34, 47
Percent female	38%	47%	51%
Smartphone usage: one week, per-person averages	**China**	**Myanmar**	**United States**
Individual's number of screenshots	29,620	11,297	16,309
Individual's total hours	40.6	15.7	22.7
Individual's number of sessions	3400	1284	2340
Individual's mean gap duration, seconds	201	400	281
Individual's mean session duration, seconds	73	47	42
Individual's session rapidity score	1.44	1.40	1.49

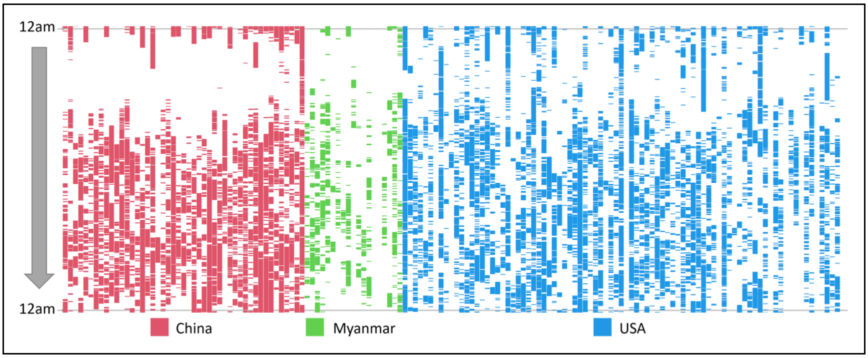

Figure 1. Visualization of subjects' smartphone screen activity throughout randomly selected 24-h periods, starting and ending at midnight. Each vertical column represents a subject; white space represents screen inactivity.

Structural variables

Based on our theory, we use five temporal variables derived from the longitudinal record of our data, summarized in Table 1, as well as two variables derived from image entropy. These variables capture a range of behaviors, reflecting varying levels of influence from contextually determined parameters versus the standardized interface. See Supplemental Appendix B for mathematical formulae.

Total hours of screentime. Overall screen time is one measure of the prevalence of the smartphone in the user's life. In our data, screenshots are captured at 5-s intervals while the screen is active. We report total screen time in hours (in one week) for interpretability.

Number of sessions. The *frequency* of smartphone use can vary even with fixed total hours of usage (Oulasvirta et al., 2012). We measure frequency via the total number of *sessions* of smartphone usage. Following previous research, a single session is defined as a period of continuous screen activity (Oulasvirta et al., 2012; Zhu et al., 2018) bounded by any gap in the 5-s interval.

Mean gap duration. This measures how long an individual typically *refrains* from activating the smartphone screen, on average (Peng et al., 2020; van Berkel et al., 2016). For each subject, this is computed by subtracting total screen time from the 168-h week, and dividing by the total number of sessions in the week.

Mean session duration. We compute each subject's average session duration by dividing total screen time by the number of sessions over the week, and converting into seconds. This metric captures how many seconds the smartphone screen holds the user's attention, on average.

Session rapidity. Most sessions are incredibly short, often under 30 s long (Reeves et al., 2021; van Berkel et al., 2016). The tendency for short sessions is captured using a power-law curve (Zhu et al., 2018). Specifically, a power-law curve is fit to each individual's histogram of session durations, and we use the best-fitting exponent as a "rapidity score." Higher levels of rapidity indicate that the distribution of sessions is more tightly clustered at the shortest durations.

Image entropy and session depth. To measure visual structure, we compute the image entropy of each screenshot; see Figure 2 for an illustration of image entropy in screenshots. We then create two subject-level image entropy metrics based on the session *depth* of each screenshot, which is the time spent "thus far" in a single smartphone usage session. The first metric is the *average image entropy of screenshots captured early in continuous smartphone usage*, which we operationalize as no more than 30 s into continuous smartphone usage. The second is the *average image entropy of screenshots captured later in continuous smartphone usage*. Thirty seconds is less than the mean session length in our data, but it is also a sufficient duration for a user to conduct habitual behaviors guided by the user interface.

For each of the seven metrics, two statistical tests are conducted to examine differences across the three samples. We use a three-way analysis of variance (ANOVA) to test against the null that the three samples have the same mean. The second test is a bootstrap-based Anderson-Darling (AD) test, which tests against the null that the three samples are drawn from the same distribution (Scholz & Stephens, 1987; Stephens, 1974). Because the sample is fairly small and unmatched,[3] we use subsetting to ensure robustness against the particular characteristics of our subjects. Each statistical test is recomputed using 12 subsets of data, according to 3 separate conditions: outlier trimming, age bracketing, and gender separation. Outliers are any subject >3 standard deviations from the sample mean on any metric (separately for our 5 temporal and 2 visual metrics). The aim of this strategy is to demonstrate consistency in findings in the face of varied approaches to analysis rather than to deduce the specific outcome of each specific test. Results from ordinary least squares (OLS) regressions of each temporal metric

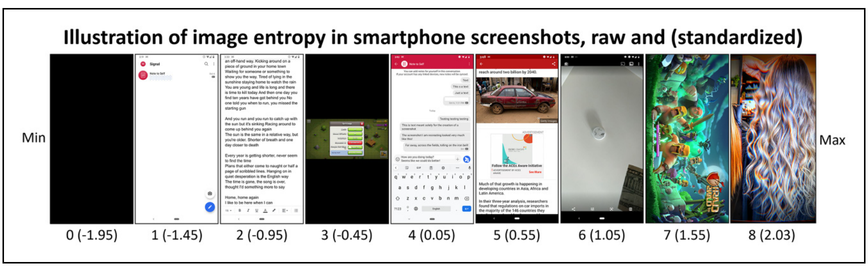

Figure 2. Image entropy in screenshots; standardization is within all screenshots in our data. For privacy, these images were created by the authors, corresponding closely to genuine images in the data.

on subsample origin, demographic features, and interactions are available in Supplemental Appendix C.

Experiential categories as content variable

Materials on the smartphone screen can be sorted into meaningful categories, which can provide insight into cross-cultural comparison on content (Goggin, 2010; Reeves et al., 2021). To build on existing research without biasing our findings in a particular direction, we compare mobile behaviors in different contexts by clustering content into broad experiential categories. Learning content features through establishing experiential categories given the massive scale of our dataset renders human annotation of on-screen content or behavior infeasible. To overcome this constraint, we combine unsupervised image clustering and human validation to derive and compare experiential content categories—that is, content that can reasonably be found in most smartphone users' records, and identified based on visual structure.

Unsupervised learning has been increasingly employed in visual analysis, as it can uncover visual patterns without predefined categories (Guérin et al., 2017; Peng T. Q., 2021). Following Peng Y (2021), we first applied transfer learning to extract visual features of all screenshots within each subsample, using the VGG16-hybrid1365 pre-trained deep learning network (Simonyan & Zisserman, 2014). This procedure outputs a feature array for each screenshot, which is a high-dimension coordinate vector capturing hundreds of values describing the composition of the screenshot.[4] We then used standard K-means clustering to group the feature arrays. To derive and compare content categories that are commensurable across subsamples, we took a two-step validation approach. First, we ran the model with K clusters, setting $K = 10$, 15, 20, 25, 30, 35, and 40. For each K-clustering solution, we randomly sampled 20 images from each cluster, and tagged these images by hand by describing what exactly was on-screen. After this first pass, we identified 10 categories that were internally coherent and mutually distinct: (a) social media usage; (b) mobile gaming; (c) music; (d) websites; (e) typing (i.e., producing text by having the keyboard shown on-screen, generally in chats); (f) full-screen image (as in watching a video or viewing a photograph); (g) full-screen text content (e.g., reading e-book); (h) maps; (i) short message service (SMS)/calls/emails; (j) phone setting screens (e.g., lock screens, home screens, and transitional screens). We then relabeled sampled screenshots of each cluster in each K-clustering solution into these 10 categories.

As our goal was to most accurately classify each of the 2.9 million screenshots into one of these 10 categories, we next determined the accuracy of each K-solution. Accuracy was measured according to the internal consistency of clusters. If our sampled clusters are internally consistent, we can feel confident that automated labeling of screenshots is reasonable and reliable. To do this, for every cluster sample of 20 screenshots, we manually calculated the proportion out of 20 that fell into one majority category. For example, if a cluster contains 16 out of 20 screenshots portraying social media behaviors, the majority category should be "social media" and the score will be 80% (16/20). This generated a score of up to 100% (20 out of 20 randomly sampled screenshots belonging to a single category). For each K-solution, we averaged these per-cluster scores across clusters. We then compared *K*-solutions according to this average score. For each country subsample, the highest-scored K-solution determined the *K* number of clusters to ultimately use in labeling all screenshots from that subsample.[5] This whole process resulted in 35 clusters for the Chinese subsample, 40 for the U.S. subsample, and 25 for the Myanmar subsample.[6] Finally, we applied the content category labeling each cluster to all screenshots assigned to that cluster in each country subsample.

Results

Localization in temporal and visual structural factors

According to Hypotheses 1a to 1c, three temporal metrics are expected to reflect localization differences across subsamples: total screen time, number of sessions, and average gap duration. Figure 3 plots total hours of screen time. The left panel of Figure 3 plots the mean and variance, while the right panel shows the entire distribution of each country sample, including subsets of the data as robustness checks. Samples are coded by color: the sample drawn from China in red, the sample drawn from the United States in blue, and the sample drawn from Myanmar in green. The left panel of Figure 3 shows that the mean hours of smartphone usage is highest in the sample drawn from China, averaging 40.6 h in the whole dataset, followed by the sample drawn from the United States, averaging 21.5 h in the whole dataset, and is lowest in the sample drawn from Myanmar, averaging 15.7 h. Across all subsets of data, the sample drawn from China shows the greatest screen time, followed by subjects in the

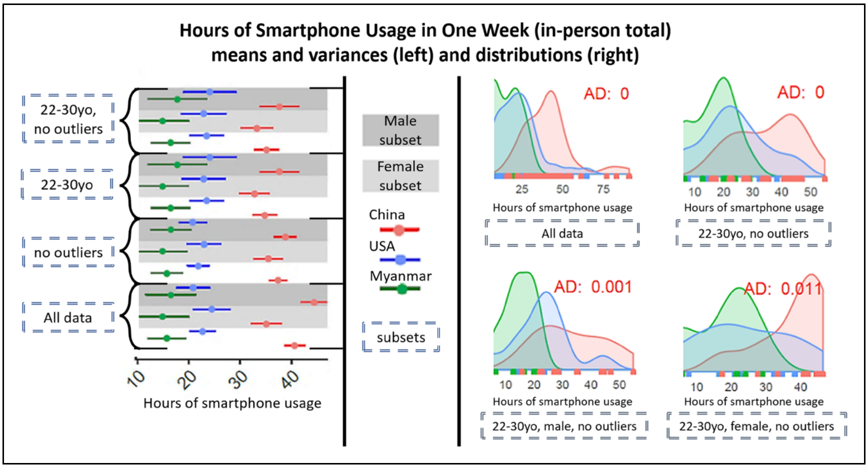

Figure 3. Total hours of smartphone usage over one week, compared across locations in multiple subsets. *p*-values for Anderson-Darling tests of distributional differences are shown in red text (≤0.05) or black text (>0.05).

United States, and then by subjects sampled from Myanmar. The means and variances differ because of substantial differences in distribution (right panel). The majority of subjects sampled from Myanmar were on their phones for less than 30 h per week, while numerous subjects sampled in China exceeded 50 h per week.

Figure 4 shows a similar pattern for the number of sessions. The sample drawn from China exhibits the highest average number of usage sessions at 3400, followed by the sample from the United States at 2178, and then the sample from Myanmar at 1284 (left panel). The sample from China exhibits the most sessions, across all subsets. Each sample's distribution of session counts is multimodal (right panel), but the sample drawn from China exhibits a consistently higher centroid than that of the sample drawn from Myanmar or the United States.

As for average gap durations, Figure 5 shows that the sample drawn from Myanmar averages about 7 min (440 s), versus about 5 min in the United States (307 s), and less than 4 min in China (201 s) using the full dataset. The sample drawn from Myanmar consistently exhibits the highest average gap duration across subsets (left panel). Notably, the typical amount of time that a subject in the sample drawn from Myanmar refrains from activating the phone screen is more than twice as long as that for a subject in the sample drawn from China. The maximum average gap duration in the entire sample is only 41 min for a member of the sample from China. Most subjects in the sample from China have an average gap duration of under 2 min (right panel). The sample from China thus has a much lower variance in average gap duration than do the samples from Myanmar and the United States.

As for visual structural features, Hypothesis 1d predicted that subjects' image entropy occurring later in usage sessions would be localized. The results generally support this

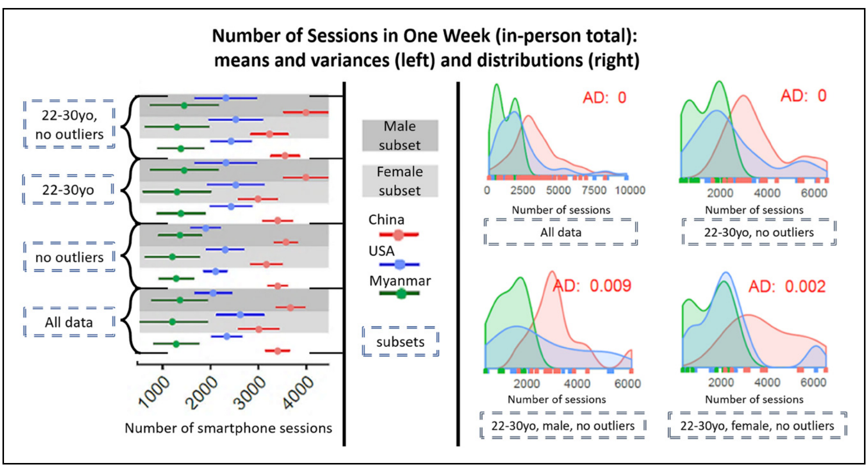

Figure 4. Total number of smartphone usage sessions over one week, compared across locations in multiple subsets. p-values for Anderson-Darling tests of distributional differences are shown in red text ($\leq$0.05) or black text (>0.05).

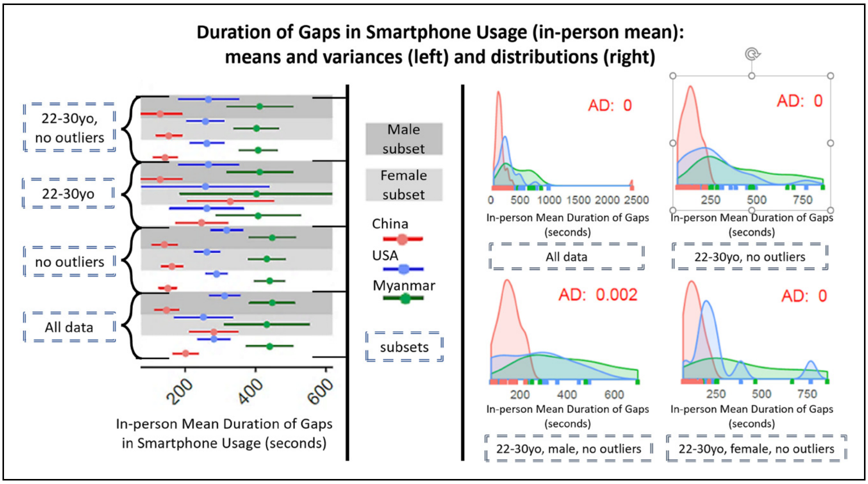

Figure 5. In-person average duration of gaps in smartphone usage, in seconds, over one week, compared across locations in multiple subsets. *p*-values for Anderson-Darling tests of distributional differences are shown in red text (≤0.05) or black text (>0.05).

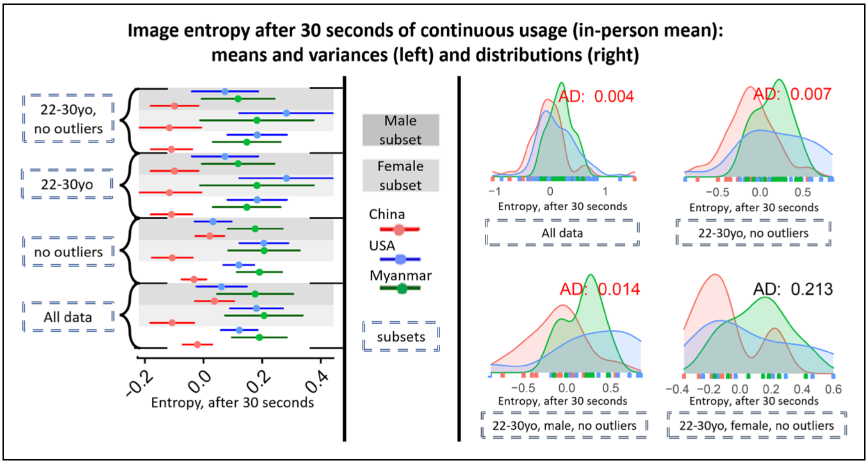

Figure 6. Image entropy computed over a week of screenshots captured after the first 30 s of a session, compared across locations in multiple subsets. *p*-values for Anderson-Darling tests of distributional differences are shown in red text (≤0.05) or black text (>0.05).

hypothesis. In Figure 6, we perform the statistical tests on subjects' average image entropy of screenshots captured *deeper* than 30 s into sessions. We find that subjects' image entropy later in the session is statistically significantly different across most subsamples, in line with the expectation that users sort into localized apps and behaviors, with differing levels of image entropy.

Homogenization: similarity in average session duration, session rapidity, and image entropy earlier in usage sessions

Hypotheses 3a and 3b predicted a lack of difference in average session duration and rapidity. The results generally support these hypotheses. Average session durations are not significantly different across samples at the 5% threshold, as shown in Figure 7. Specifically, the average session duration falls between 40 s and 50 s across virtually all further subsets. Three notable outliers are in the sample drawn from China, but differences in the three samples are not significant when using the full sample, nor when using any subsets thereof (right panel). With outliers removed, average session durations are similarly distributed between 20 s and 80 s across the 3 subsamples.

Figure 8 shows that rapidity scores are not significantly different across the three subsamples at the 5% threshold. Rapidity scores fall between 1.4 and 1.6 across subsets. This indicates that phone usage most commonly occurs in bursts of short sessions, with rare instances of sustained usage (left panel). The sample drawn from the United States exhibits a rightward skew; this is also evident in the mean differences of the age subsets in the left panel. These higher rapidity scores among subjects in the sample drawn from the United States do not meet the outlier criteria, and they represent a pattern of usage with virtually no sessions longer than a couple of

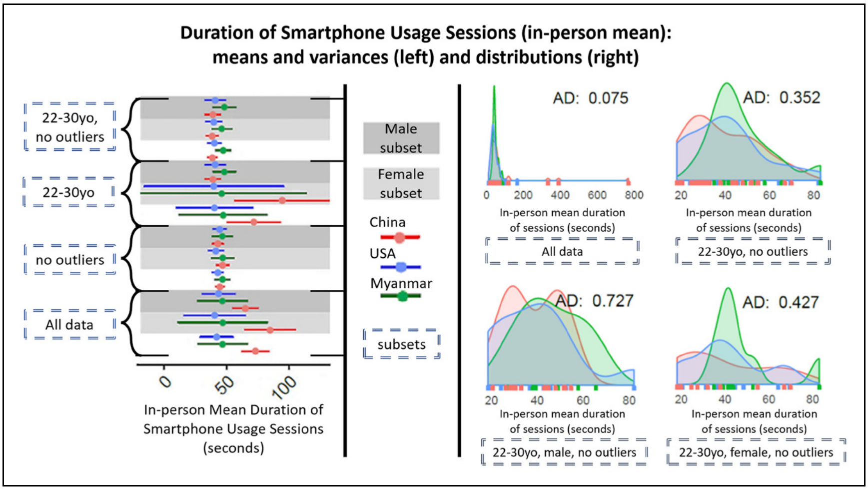

Figure 7. Average session duration over one week, in seconds, compared across locations in multiple subsets. *p*-values for Anderson-Darling tests of distributional differences are shown in red text (≤0.05) or black text (>0.05).

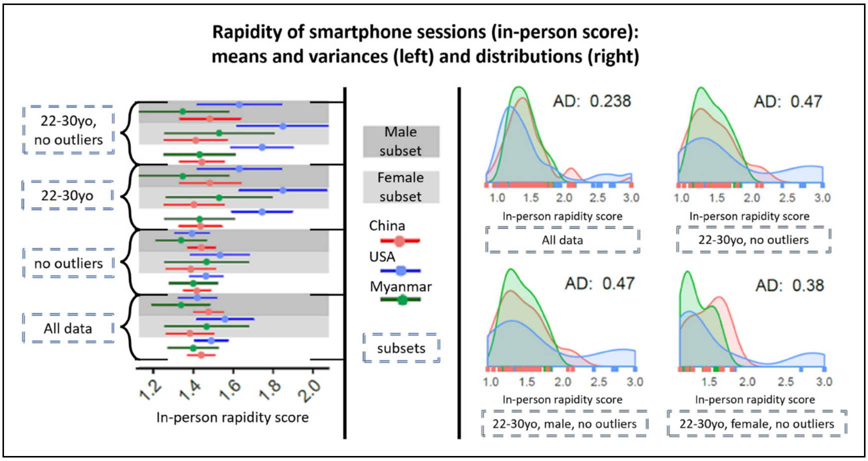

Figure 8. Session rapidity scores computed over a week, compared across locations in multiple subsets. *p*-values for Anderson-Darling tests of distributional differences are shown in red text (≤0.05) or black text (>0.05).

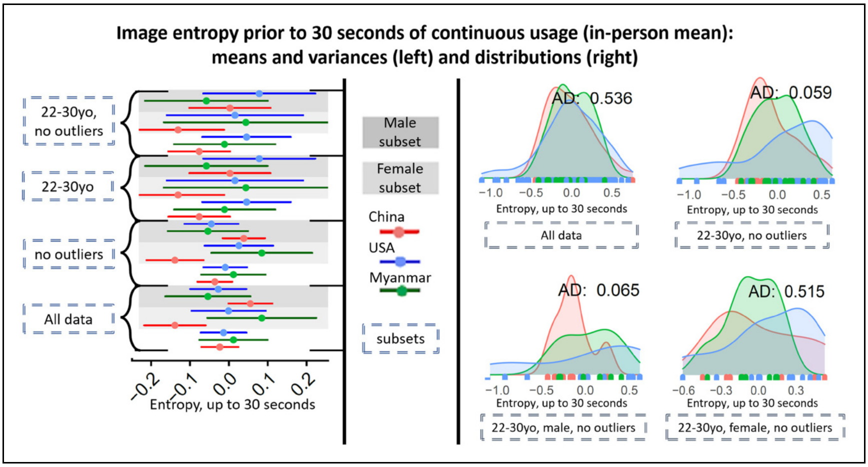

Figure 9. Image entropy computed over a week in screenshots captured in the first 30 s of a session, compared across locations in multiple subsets. *p*-values for Anderson-Darling tests of distributional differences are shown in red text (≤0.05) or black text (>0.05).

minutes, despite frequent smartphone use. In contrast, the lower rapidity scores indicate a broader intra-individual distribution of session durations, often lasting longer than a couple of minutes.

Hypothesis 3c predicts that image entropy earlier in usage sessions would be homogeneous. The results generally support this hypothesis. As shown in Figure 9, image entropy before 30 s is not significantly different across samples at the 5% threshold. In other words, it shows the homogeneity in habitual mobile phone usage. Each subsample has a wide range of typical image entropy at shallow session depth.

Homogenization and localization in experiential content categories

Table 2 shows the average proportion of screenshots among users in each subsample on 10 on-screen behaviors, along with p-value results from one-way ANOVA testing. We find significant differences across subsamples in seven behaviors and no such differences in three behaviors. Within these 10 categories, we focus on the 2 expectations of our Hypotheses 2 and 4. These two categories are well represented across subsamples, owing to their definitional broadness and reasonably universal utility, and present an interesting highlight of homogenization and localization.

While ample social media use is common in all three subsamples, a full 52.1% of screenshots in Myanmar were classified as social media, reflecting Facebook's well-known dominance in that media environment (Leong, 2017). Likewise, 41.2% of screenshots in China were classified as social media, reflecting the dominance of social media like WeChat and Weibo in China. Only a quarter of smartphone screenshots in the United States were taken up by social media. In contrast, typing is very homogeneous across groups. In the United States, China, and Myanmar subsamples, the keyboard is present on screen for roughly 14% of total screenshots, and is the most homogeneous behavior out of all those detected by our algorithm.

Discussion and conclusion

Through analyzing moment-by-moment screenshot log data, this study shows that localization and homogenization are *both* reflected in smartphone usage. Across our three distinct groups, we find significant differences in multiple structural measures of overall usage, which falls in line with expectations of localization according to contextually

Table 2. Average proportions of screenshots in each of 10 categories, per national subsample, with p-values from one-way analysis of variance (ANOVA) testing.

Category	China	Myanmar	United States	ANOVA p-value
Full screen image	0.13	0.08	0.09	0.08
Full screen text	0.14	0.00	0.00	0.00***
Keyboard	0.12	0.14	0.13	0.82
Maps	0.00	0.00	0.04	0.00***
Mobile gaming	0.04	0.04	0.11	0.001***
Music	0.01	0.00	0.00	0.08
Phone setting screens	0.13	0.19	0.19	0.00**
SMS/calls/emails	0.00	0.05	0.09	0.00***
Social media usage	0.43	0.52	0.24	0.00***
Websites	0.00	0.00	0.10	0.00***

*$p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

determined parameters. In contrast, we find homogeneity in smartphone behaviors comparatively more driven by the standardized interface, focused on the speed and pacing of smartphone usage sessions. Likewise, the visual structure of smartphone usage is distinct across groups only after an initial period of presumably routine screen interaction. We also find that user location moderates the degree to which social media dominates smartphone usage (as well as behaviors like mobile gaming and map usage), which follows logically from known characteristics of the three mobile ecosystems. This contrasts with interestingly homogeneous patterns of text production across subsamples. Altogether, multiple measurement styles have identified the same point: homogenization and localization are simultaneously present in smartphone usage.

We contribute a theoretical distinction in identifying the variable that determines whether localization or homogenization will dominate in a particular aspect of smartphone usage: the relevance of contextually determined parameters, relative to the interface and the behaviors it elicits. Substantively, the structural similarity we find in session durations and other granular behaviors adds a new layer of support to scholars critical of homogenization and its potential effects on global citizens. Mansell (2017) and others describe the imposition of homogeneous American values through the spread of new media. Insofar as new media structure the psychological and social experience of time (Kaun et al., 2016), there may be a global standardization of the experience of time that results from the use of a standard device. Future research could directly investigate the extent to which user interface design and temporal processing interact with one another. In particular, longer-lasting samples could better measure the slow dynamic processes of homogenization and localization, building on the general description we offer here.

This paper also makes methodological contributions to mobile studies. The data collection method and sample used in this study offer possibilities for future empirical social science research regarding the smartphone. The Screenomics framework complements survey research with granular behavioral data that cannot easily be self-reported. Future research can leverage this framework to understand mobile phone usage of people in more diverse contexts and social status. The synthesis of unsupervised learning on visual elements, transfer learning, and human validation also represents the potential of employing computational methods and the importance of the image-as-data framework (Joo et al., 2018; Zhang & Pan, 2019) in understanding mobile behaviors. More studies in mobile communication can employ these approaches to derive user behaviors from mobile image data and improve these methods to fit in different research scenarios.

As smartphone-logging technology becomes more accessible to researchers, observational data collection is increasingly scalable, widening the scope of future research. Future research would employ larger cross-sectional samples without sacrificing the richness afforded by screen recording, as this would potentially build more accurate representations of cultures and subcultures. The contextually determined parameters that underlie our hypotheses were left unenumerated in our study, and may be rooted in many different demographic or environmental variations present across our three samples, or their interactions. We hope the methods and results of this study provide a starting point for future research to capture a more complete picture of the moderating factors in the structure of device usage.

Acknowledgments

Our thanks to participants of the Human Screenome Project in the United States, China, and Myanmar; Haoran Li, Xiuyu Li, Mingqi Song, and Jason Yar for research assistance; Nilam Ram and Thomas Robinson for data and research support, and the following members of the Screenomics Lab: Miriam Brinberg, Agnese Chiatti, Mu-Jung Cho, Andrew Fitzgerald, Anupriya Gagneja, Jihye Lee, Katherine Roehrick, and Xiao Yang; Seth Geiger at SmithGeiger Group; Jonathan Zhu, Winson Peng, Jenny Ju, and participants of the International Communication Association (ICA) 68th Annual Conference (Information Systems Division), the Stanford Data Science Lab, Conversation Series in the United Nations University Institute in Macau, 2021 China Computational Social Science Forum, and 2021 ICA Regional Hub Symposium at Central South University for many helpful comments and suggestions.

Funding

We thank the Cyber Social Initiative at Stanford University [SPO#125124], the Stanford University King Center on Global Development, the Knight Foundation [G-2017-54227], the National Science Foundation (NSF) (Grant #1831481), and the Stanford Graduate Fellowship for funding assistance.

ORCID iD

Yingdan Lu https://orcid.org/0000-0002-9955-6070

Supplemental material

Supplemental material for this article is available online.

Notes

1. In other research, image entropy has been linked to engagement, attention, recall, and even perception of time (Palumbo et al., 2014; Tuch et al., 2009).
2. A modification was made for Chinese smartphones, which avoided using any banned Google services (see Supplemental Appendix A).
3. Matching on age and gender resulted in non-convergence; other demographic variables are incommensurable.
4. As the feature output of each image yields 4,096 dimensions, to facilitate faster computation, we de-dimensionalized each feature vector using the principal component analysis, and kept the first 481 factors, which retains more than 80% of the variance for data in each country.
5. Note that one content category can dominate multiple clusters.
6. See Supplemental Appendix D for the average (Yang et al., 2019) cluster-consistency scores for optimal *K*s.

References

Bagozzi, R. P. (2007). The legacy of the technology acceptance model and a proposal for a paradigm shift. *Journal of the Association for Information Systems*, *8*(4), 3. 10.17705/1jais.00122

Bagozzi, R. P., Davis, F. D., & Warshaw, P. R. (1992). Development and test of a theory of technological learning and usage. *Human Relations*, *45*(7), 659–686. 10.1177/001872679204500702

Baron, N. S., & Segerstad, Y. H. (2010). Cross-cultural patterns in mobile-phone use: Public space and reachability in Sweden, the USA and Japan. *New Media & Society*, *12*(1), 13–34. 10.1177/1461444809355111

Boase, J., & Ling, R. (2013). Measuring mobile phone use: Self-report versus log data. *Journal of Computer-Mediated Communication*, *18*(4), 508–519. 10.1111/jcc4.12021

Brinberg, M., Ram, N., Yang, X., Cho, M. J., Sundar, S. S., Robinson, T. N., & Reeves, B. (2020). The idiosyncrasies of everyday digital lives: Using the human screenome project to study user behavior on smartphones. *Computers in Human Behavior*, *114*(106570), 106570. 10.1016/j.chb.2020.106570

Campbell, S. W. (2007). A cross-cultural comparison of perceptions and uses of mobile telephony. *New Media & Society*, *9*(2), 343–363. 10.1177/1461444807075016

Cho, M. J. (2020). *Media Production Behavior on Smartphones: Time, Content, Structure, and Context*.

Goggin, G. (2010). *Global mobile media*. Routledge.

Guérin, J., Gibaru, O., Thiery, S., & Nyiri, E. (2017). CNN Features are also great at unsupervised classification. arXiv preprint arXiv:1707.01700

Harari, G. M., Lane, N. D., Wang, R., Crosier, B. S., Campbell, A. T., & Gosling, S. D. (2016). Using smartphones to collect behavioral data in psychological science: Opportunities, practical considerations, and challenges. *Perspectives on Psychological Science*, *11*(6), 838–854. 10.1177/1745691616650285

Hoskins, C., & Mirus, R. (1988). Reasons for the US dominance of the international trade in television programmes. *Media, Culture and Society*, *10*(4), 499–515.

Humphreys, L., Karnowski, V., & von Pape, T. (2018). Smartphones as metamedia: A framework for identifying the niches structuring smartphone use. *International Journal of Communication*, *12*, 17. 1932–8036/20180005

International Telecommunication Union (n.d.). Mobile cellular subscriptions (per 100 people). In World Telecommunication/ICT Development Report and database. Retrieved March 27, 2019, from https://data.worldbank.org/indicator/IT.CEL.SETS.P2?end=2017&locations=MM

Jenkins, H. (2004). The cultural logic of media convergence. *International Journal of Cultural Studies*, *7*(1), 33–43. 10.1177/1367877904040603

Jensen, K. B. (2011). Meta-media and meta-communication—revisiting the concept of genre in the digital media environment. *MedieKultur: Journal of Media and Communication Research*, *27*(51), 14. 10.7146/mediekultur.v27i51.4032

Joo, J., Steinert-Threlkeld, Z. C., & Luo, J. (2018-10). Social and political event analysis based on rich media. *Proceedings of the 26th ACM international conference on multimedia*, 2093–2095.

Kang, H., & Sundar, S. S. (2016). When self is the source: Effects of media customization on message processing. *Media Psychology*, *19*(4), 561–588. 10.1080/15213269.2015.1121829

Kappos, A., & Rivard, S. (2008). A three-perspective model of culture, information systems, and their development and use. *MIS Quarterly*, *32*(3), 601–634. 10.2307/25148858

Kaun, A., Fornäs, J., & Ericson, S. (2016). Media times: mediating time—temporalizing media: Introduction. *International Journal of Communication*, *10*(7). 1932–8036/20160005

Kay, A., & Goldberg, A. (1977). Personal dynamic media. *Computer*, *10*(3), 31–41. 10.1109/C-M.1977.217672

Keller, W. (2002). Geographic localization of international technology diffusion. *American Economic Review*, *92*(1), 120–142. 10.1257/000282802760015630

King, G., Pan, J., & Roberts, M. E. (2013). How censorship in China allows government criticism but silences collective expression. *American Political Science Review*, *107*(2), 326–343. :10.1017/S0003055413000014

King, G., Pan, J., & Roberts, M. E. (2014). Reverse-engineering censorship in China: Randomized experimentation and participant observation. *Science* (6199), 891–891. 10.1126/science.1251722

Kobayashi, T., & Boase, J. (2012). No such effect? The implications of measurement error in self-report measures of mobile communication use. *Communication Methods and Measures*, *6*(2), 126–143. 10.1080/19312458.2012.679243

Langley, P., & Leyshon, A. (2017). Platform capitalism: The intermediation and capitalisation of digital economic circulation. *Finance and Society*, *3*(1), 11–31. 10.2218/finsoc.v3i1.1936

Leong, L. (2017). Mobile Myanmar: The development of a mobile app culture in Yangon. *Mobile Media & Communication*, *5*(2), 139–160. 10.1177/2050157917689950

Lin, Y. H., Lin, Y. C., Lee, Y. H., Lin, P. H., Lin, S. H., Chang, L. R., Tseng, H. W., Yen, L. Y., Yang, C. H., & Kuo, B. J. (2015). Time distortion associated with smartphone addiction: Identifying smartphone addiction via a mobile application (App). *Journal of psychiatric research*, *65*, 139–145. 10.1016/j.jpsychires.2015.04.003

Ling, R., Panchapakesan, C., Aricat, R., Oreglia, E., & Lwin, M. O. (2017). Mobile communication in Myanmar. In Goggin, G., & McLelland, M. (Eds.), *The Routledge companion to global internet histories* (pp. 296–310). Routledge.

Mansell, R. (2012). *Imagining the internet: Communication, innovation, and governance*. Oxford University Press.

Mansell, R. (2017). Imaginaries, values and trajectories: A critical reflection on the internet. In Goggin, G., & McLelland, M. (Eds.), *The Routledge companion to global internet histories* (pp. 23–33). Routledge.

Mante, E. (2002). The Netherlands and the USA compared. In Katz and, J., & Aakhus, M. (Eds.), *Perpetual contact: Mobile communication, private talk, public performance* (pp. 110–125). Cambridge University Press.

Mehrotra, A., Müller, S. R., Harari, G. M., Gosling, S. D., Mascolo, C., Musolesi, M., & Rentfrow, P. J. (2017). Understanding the role of places and activities on mobile phone interaction and usage patterns. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *1*(3), 84. 10.1145/3131901

Montag, C., Becker, B., & Gan, C. (2018). The multipurpose application WeChat: A review on recent research. *Frontiers in Psychology*, *9*. 10.3389/fpsyg.2018.02247

Oulasvirta, A., Rattenbury, T., Ma, L., & Raita, E. (2012). Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing*, *16*(1), 105–114. 10.1007/s00779-011-0412-2

Oulasvirta, A., Tamminen, S., Roto, V., & Kuorelahti, J. (2005, April). Interaction in 4-s bursts: The fragmented nature of attentional resources in mobile HCI. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 919–928). ACM.

Palumbo, L., Ogden, R., Makin, A. D., & Bertamini, M. (2014). Examining visual complexity and its influence on perceived duration. *Journal of Vision*, *14*, 3–3. 10.1167/14.14.3

Peng, T. Q., Zhou, Y., & Zhu, J. J. (2020). From filled to empty time intervals: Quantifying online behaviors with digital traces. *Communication Methods and Measures*, *14*(4), 1–20. 10.1080/19312458.2020.1812556

Peng, Y. (2021). What makes politicians' Instagram posts popular? Analyzing social media strategies of candidates and office holders with computer vision. *International Journal of Press/Politics*, *26*(1), 143–166. 10.1177/1940161220964769

Postill, J. (2008). Localizing the internet beyond communities and networks. *New Media & Society*, *10*(3), 413–431. 10.1177/1461444808089416

Price, M. E. (2002). *Media and sovereignty: The global information revolution and its challenge to state power*. MIT press.

Punathambekar, A., & Mohan, S. (2019). Digital platforms, globalization and culture. In Curran, James, & & Hesmondhalgh, David (Eds.), *Media and Society* (6th, pp. 207–226).

Qin, X., Tan, C. W., Law, E. L. C., Bødker, M., Clemmensen, T., Qu, H., & Chen, D. (2018, April). Deciphering the role of context in shaping mobile phone usage: Design recommendations for context-aware mobile services from a cross-cultural perspective. In *Proceedings of the Sixth International Symposium of Chinese CHI* (pp. 39-48). ACM.

Reeves, B., Ram, N., Robinson, T. N., Cummings, J. J., Giles, C. L., Pan, J., Chiatti, A., Cho, M. J., Roehrick, K., Yang, X., Gagneja, A., Brinberg, M., Muise, D., Lu, Y., Luo, M., Fitzgerald, A., & Yeykelis, L. (2021). Screenomics: A framework to capture and analyze personal life experiences and the ways that technology shapes them. *Human–Computer Interaction*, *36*(2), 150–201. 10.1080/07370024.2019.1578652

Reeves, B., Robinson, T. N., & Ram, N. (2020). Time for the human screenome project. *Nature*, *577*, 314–317. 10.1038/d41586-020-00032-5

Reeves, B., Thorson, E., & Schleuder, J. (1985). Attention to television: psychological theories and chronometric measures. *Perspectives on Media Effects*, *27*(3-4), 251–279. 10.3109/00207458509149770

Rothschild, M. L., Thorson, E., Reeves, B., Hirsch, J. E., & Goldstein, R. (1986). EEG Activity and the processing of television commercials. *Communication Research*, *13*(2), 182–220. 10.1177/009365086013002003

Schiller, H. (1976). *Communication and cultural domination*. ME Sharpe.

Scholz, F. W., & Stephens, M. A. (1987). K-sample Anderson–Darling tests. *Journal of the American Statistical Association*, *82*(399), 918–924. 10.1.1.1087.8833

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423. 10.1002/j.1538-7305.1948.tb01338.x

Shepard, C., Rahmati, A., Tossell, C., Zhong, L., & Kortum, P. (2011). LiveLab: measuring wireless networks and smartphone users in the field. *ACM SIGMETRICS Performance Evaluation Review*, *38*(3), 15–20.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.*1556 arXiv:1409.1556

Smith, C. L., & Kantor, P. B. (2008, July). User adaptation: Good results from poor systems. In *Proceedings of the 31st annual international ACM SIGIR conference on research and development in information retrieval* (pp. 147–154). *Association for Computing Machinery* .

Sparks, C. (2012). Media and cultural imperialism reconsidered. *Chinese Journal of Communication*, *5*(3), 281–299. 10.1080/17544750.2012.701417

Statista (2019a). Smartphone penetration rate in China from 2015 to 2023. Retrieved March 27, 2019, from https://www.statista.com/statistics/321482/smartphone-user-penetration-in-china/

Statista (2019b). Smartphone user penetration as share of population in the United States from 2017 to 2023. Retrieved March 27, 2019, from https://www.statista.com/statistics/201184/percentage-of-mobile-phone-users-who-use-a-smartphone-in-the-us/

Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, *69*(347), 730–737. 2286009

Sundar, S. S., & Oh, J. (2019). Psychological effects of interactive media technologies: A human–computer interaction (HCI) perspective. In M. B. Oliver, A. A. Raney, & J. Bryant (Eds.), *Media Effects: Advances in Theory and Research* (4th ed., pp. 357–372). Routledge.

Taneja, H., & Wu, A. X. (2014). Does the great firewall really isolate the Chinese? Integrating access blockage with cultural factors to explain web user behavior. *Information Society*, *30*(5), 297–309. 10.1080/01972243.2014.944728

Tuch, A. N., Bargas-Avila, J. A., Opwis, K., & Wilhelm, F. H. (2009). Visual complexity of websites: effects on users' experience, physiology, performance, and memory. *International Journal of Human–Computer Studies*, *67*(9), 703–715. 10.1016/j.ijhcs.2009.04.002

Turow, J. (2011). *Media today: An introduction to mass communication*. Taylor & Francis.

Van Alstyne, M., & Brynjolfsson, E. (2005). Global village or cyber-Balkans? Modeling and measuring the integration of electronic communities. *Management Science*, *51*(6), 851–868. 10.1287/mnsc.1050.0363

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

West, S. (2019). Data capitalism: Redefining the logics of surveillance and privacy. *Business and society*, *58*(1), 20–41. 10.1177/0007650317718185

Wu, A. X., & Taneja, H. (2016). Reimagining internet geographies: A user-centric ethnological mapping of the world wide Web. *Journal of Computer-Mediated Communication*, *21*(3), 230–246. 10.1111/jcc4.12157

Yang, X., Ram, N., Robinson, T., & Reeves, B. (2019). Using screenshots to predict task switching on smartphones. *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–6.

Yeykelis, L., Cummings, J. J., & 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. 10.1111/jcom.12070

Yu, H. (2017). Social imaginaries of the internet in China. In Goggin, G., & McLelland, M. (Eds.), *The Routledge companion to global internet histories* (pp. 296–309). Routledge.

Zhang, H., & Pan, J. (2019). CASM: A Deep-Learning Approach for Identifying Collective Action Events with Text and Image Data from Social Media. *Sociological Methodology*, *49*(1), 1–57. http://dx.doi.org/10.1177/0081175019860244

Zheng, X. S., Chakraborty, I., Lin, J. J. W., & Rauschenberger, R. (2009, April). Correlating low-level image statistics with users-rapid aesthetic and affective judgments of web pages. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1-10). *Association for Computing Machinery*.

Zhu, J. J., Chen, H., Peng, T. Q., Liu, X. F., & Dai, H. (2018). How to measure sessions of mobile phone use? Quantification, evaluation, and applications. *Mobile Media & Communication*, *6*(2), 215–232. 10.1177/2050157917748351

Author Biographies

Daniel Muise is a PhD Candidate in the Department of Communication at Stanford University. His research focuses on the interaction of time, platform design, and media psychology in the consumption of information, and its relation to system-level processes in political communication. He applies computational social science tools to study temporal assumptions underlying mass polarization, engagement with news, and methodological challenges to user experience research.

Yingdan Lu is a PhD Candidate in the Department of Communication, Stanford University. Her research focuses on digital media, political communication, and online information manipulation. She employs computational methods on massive multimodal datasets to understand social media propaganda, media governance, misinformation and disinformation, and the presence of digital inequality around the world and on diverse platforms. Her research has been published in peer-

reviewed journals, including *Political Communication*, *Human-Computer Interaction*, and among other peer-reviewed journals.

Jennifer Pan (PhD, Harvard University) is an Associate Professor of Communication at Stanford University. Her research focuses on political communication and authoritarian politics. She uses experimental and computational methods with large-scale datasets on political activity in China and other authoritarian regimes to answer questions about how autocrats perpetuate their rule; how political censorship, propaganda, and information manipulation work in the digital age; and how preferences and behaviors are shaped as a result.

Byron Reeves (PhD, Michigan State University) is the Paul C. Edwards Professor of Communication at Stanford University, and Professor (by courtesy) in the Stanford School of Education. His research is about the psychological processing of media, and resulting responses and effects. He has studied how media influence attention, memory and emotional responses and has applied the research in the areas of speech dialogue systems, interactive games, advanced displays, social robots, and autonomous cars.