How consistent are smartphone application preferences? A descriptive study of mobile application repertoires using behavioral data

Douglas A. Parry

Department of Information Science, Stellenbosch University, South Africa

Craig J. R. Sewall

Department of Psychiatry, University of Pittsburgh, USA

Acknowledgements:

CJRS is supported by the National Institute of Mental Health (T32MH018951).

Corresponding author:

Douglas A. Parry: dougaparry@sun.ac.za

Abstract

Smartphones afford users the ability to select their own mobile application repertoires through the installation and use of various applications. We report a quantitative descriptive study of the types of applications that people commonly use, the amount of time they spend with these applications, the application combinations that they construct, the consistency of these combinations over time, and the differences in these outcomes by demographic characteristics. Using a longitudinal dataset collected from a U.S. adult sample during the COVID-19 pandemic, the study leverages behavioral data collected via data donations to identify key application adoption patterns and shows that peoples' mobile application repertoires are concentrated around a set of popular applications that is relatively consistent over time. However, within this set there is considerable diversity between individuals and applications, suggesting that quantifying smartphone usage with a single metric— 'screentime'—is unlikely to capture the full extent of media that users engage with.

Keywords: Smartphone, usage, screentime, mobile applications, digital trace data, data donation

How consistent are smartphone application preferences? A descriptive study of mobile application repertoires using behavioral data

The convergence of most modern communications media into a single metamedium—a smartphone—implies that, with few exceptions, smartphones afford users the ability to gratify almost all their needs with just a single device (Humphreys, 2018). Whether someone wants to communicate with their family members, organize an event with friends, check work or school updates, or read the news, they will invariably use a smartphone. This convergence is primarily enabled by the capacity to configure the features available on a smartphone through the installation of mobile applications. Such constituent media (Humphreys, 2018) afford users the ability to uniquely configure the action possibilities that their smartphone enables through the crafting of their own mobile application repertoires (Hasebrink & Popp, 2006; Heeter, 1985).

As a result of the diverse uses and gratifications offered by smartphones, these devices now occupy an ever-growing proportion of peoples' time (Schumacher & Kent, 2020). This has prompted researchers to focus on the effects that ubiquitous smartphone use might produce (for a review see Meier et al., 2020). While addressing such concerns is important, there remain fundamental questions about actual smartphone usage that also require attention (Festic et al., 2021; Munger et al., 2021). The dynamic and customisable nature of smartphone applications, coupled with the rapid pace with which new features are deployed, means, firstly, that our collective conception of what smartphone use actually entails requires continual updating and, secondly, that collapsing all smartphone use into a single high-level metric—screentime—will not capture the full range of uses that these devices enable. Smartphone usage is not a static construct that can be defined once and left in place, nor is it a single behavior that is the same for all users. Munger et al. (2021) acknowledge this dynamism in their call for increased attention to the description of digital media usage. Festic et al. (2021 p. 2), similarly, argue that “advancing

media and communication theories requires solid empirical evidence on internet use and a foundational understanding of the scale of use in everyday life.”

We argue that such a need also exists for smartphone use. The purpose of this study is to investigate the applications that people commonly use, the amount of time they spend with these applications, the application combinations that they construct, the consistency of these combinations over time, and the degree to which these outcomes are influenced by various demographic characteristics. To address these objectives, and overcome shortcomings in extant findings, we leveraged a longitudinal dataset collected from a U.S. adult sample during the COVID-19 pandemic that includes behavioral data on smartphone application usage.

Mobile application repertoires

First used to describe the selection of television channels from the bouquet offered by a cable provider (Ferguson, 1992; Heeter, 1985), the term “repertoire” refers to the selection of a subset of preferred options out of available choices (Ferguson, 1992; Heeter, 1985). Expanding from this video-oriented context, Hasebrink and Popp (2006), propose the broader concept of a “media repertoire” to describe the comprehensive patterns of media use that result from the continued selection and regular use of a preferred subset of media from all available choices (e.g., books, radio, internet, video, magazines, etc.). Building on these conceptions, we use the term *mobile application repertoire* to refer to the set of preferred mobile applications that an individual selects and regularly uses. This definition acknowledges that smartphones, and other mobile devices like tablets and smartwatches more broadly, are metamedia that bundle various constituent media or, in this case, mobile applications (Humphreys, 2018). Further, it also acknowledges that these metamedia are continually subjected to user-driven configuration through the installation of an evolving set of constituent media (Humphreys, 2018; Jensen,

2016). An individual's mobile application repertoire, therefore, represents the applications that they have selected to regularly use out of the nearly endless array of choices available.

Building on earlier studies primarily focusing on the selection of information sources (Kim, 2016; Wolf & Schnauber, 2015; Yuan, 2011), researchers have used the repertoire approach to investigate use of digital media like social networking services and mobile applications (Boczkowski et al., 2018; Haight et al., 2014; Horvát & Hargittai, 2021; Matassi et al., 2022; Taneja et al., 2012). Horvát and Hargittai (2021), for instance, investigated co-occurrences in social media use. In a survey of U.S. adults in 2016, respondents indicated whether they currently used 10 social media platforms specified by the researchers. Using network analyses, this study identified pairs of platforms that were commonly co-adopted. The findings also revealed that gender was a key predictor of platform adoption, with women and men choosing different combinations. Similar differences were found for age and education level. These results support earlier research among Canadians which revealed demographic differences in levels of online activity and platform usage (Haight et al., 2014). Specifically, Haight et al. (2014) found that social media use differs by age, gender, and education level. Such differences have also been reported among Argentinian internet users (Matassi et al., 2022).

Focusing on mobile applications, Griffioen et al. (2021) investigated use of smartphone applications through video observation during a 10-minute break. The investigation revealed substantial diversity in how people used their smartphones, and the authors concluded that "screen time is a term that holds little descriptive or explanatory power anymore" (p. 8). While some applications are certainly more popular than others, Griffioen et al. (2021) draw attention to the diverse combinations of applications and show that people engage with a variety of widely known (e.g., *Facebook*, *Instagram*) and more niche platforms (e.g., *Tinder*, *Quora*).

Recent studies have begun to leverage digital trace data to investigate use of specific applications or application categories (Festic et al., 2021; Ohme et al., 2021; Shaw et al., 2022). Because digital trace data represent the ‘residue’ of everyday actions that are logged as a byproduct of device usage (Freelon, 2014), they overcome many of the shortcomings of existing research that primarily relies on self-reported estimates of usage duration or frequency (Dickson et al., 2018; Griffioen et al., 2020). Meta-analytic evidence shows that there is a considerable discrepancy between participants’ estimates of their usage and metrics produced from digital trace data (Parry et al., 2021). Usage logs represent a more valid, albeit imperfect (Jürgens et al., 2019), representation of usage than retrospective estimates.

To investigate communication processes that draw on smartphone usage, a growing number of researchers have turned to data donation (Boeschoten et al., 2020; Ohme et al., 2021) and recruited participants to take screenshots to capture information on usage collected passively in the course of normal device use (Jones-Jang et al., 2020; Ohme et al., 2021; Sewall & Parry, 2021). Most studies using data donation have, however, focused on aggregate measures of usage like the overall usage duration, the number of pickups, or the number of notifications (Jones-Jang et al., 2020; Ohme et al., 2021). Such metrics only provide a high-level perspective on device usage. If we are to understand the application repertoires that users craft, and the possible implications that this behavior holds for people’s wellbeing (Büchi, 2020), we need to leverage behavioral usage metrics that capture more specific aspects of media usage (Parry et al., 2022).

Studies have used digital trace data to investigate use of a specific basket of applications selected a priori (Ohme et al., 2021; Sewall et al., 2021; Walsh et al., 2021). Some researchers (e.g., Walsh et al., 2021) have produced their own categories by selecting specific applications prior to data collection, while others (e.g., Sewall et al., 2021) have analyzed usage in categories

pre-specified by tracking applications. Kwon et al. (2021) for instance, collected usage logs among a Korean sample in 2019 to assess the efficacy of using formal concept analysis to identify mobile application repertoires. Among this sample, the analysis revealed that mobile application repertoires—specified on the basis of the categories provided by the Google App Store—differed by age, gender, and specific age-gender combinations. Festic et al. (2021), used tracking software to investigate use of six prominent services across mobile and desktop devices among a Swiss sample. While their data were collected over three months, they did not investigate the consistency of application usage over the tracking period. Instead, all results were reported in aggregate. In contrast to the general practice of using cross-sectional or aggregate usage data, Shaw et al. (2022) investigated the consistency with which people used 21 popular smartphone applications over a seven day period, finding that, over this period, people display distinctive patterns of digital behavior.

Taken together, while most studies have primarily focused on a predetermined basket of applications, relied on self-reported data, and only considered use versus non-use of a given platform, they all generally support a common conclusion: While many similarities exist, there is also diversity in application repertoires, and this diversity is, at least in part, driven by demographic factors like age, gender, or education level, among other individual differences. There remain, however, key gaps in our understanding of mobile application repertoires.

The Present Study

In this paper we report on behavioral data collected as part of a longitudinal study running from August 2020 to November 2020. This four-wave study with one month intervals includes log data from a sample of adults residing in the U.S. We used this dataset to address the following research questions:

- RQ1: (a) *Which applications are most frequently adopted, and (b) how much time do people spend using these applications?*
- RQ2: *How do peoples' application choices differ according to demographic characteristics?*
- RQ3: *How does application usage time differ according to demographic characteristics?*
- RQ4: *How consistent are peoples' application choices over time?*

The code and (anonymised) data used to address these research questions is available on the OSF (https://osf.io/4vbq9/?view_only=bc0cac0c83284b688905557aaf16fb4c).

Method

Procedures

Following approval by the University of [redacted for review] ethics review board, participants were recruited via *Prolific*, an online recruitment service targeted at the academic market. Volunteers were eligible to participate if they were: (1) U.S. residents, (2) 18-35 years old, (3) iPhone users, and (4) had ≥ 10 submissions on *Prolific* with task approval rating $\geq 95\%$. Participants were routed to a Qualtrics survey for data collection. Those who completed the survey at wave one were followed-up to participate in the next three waves. Participants were compensated \$4.00 for the first wave and \$3.00 for each subsequent wave. Additionally, those who participated in all four waves received a further \$2.00.

Participants

Three hundred ninety-six participants provided data for the first wave, of which 12 were excluded due to submitting inauthentic screenshots (i.e., they provided screenshots that were stock images), failing multiple attention checks in the survey, or failing to submit the correct *Prolific* authentication code. These exclusions resulted in a baseline sample of $n = 384$. For the second wave, which took place in September 2020, 337 participants provided authentic

screenshots of their usage. The third wave, in October 2020, included the data of 318 participants, while the fourth wave, in November 2020, included the data of 308 participants. The study, therefore, had an attrition rate of 19.79%. However, only 286 participants provided data for all four waves—a completion rate of 74.48%. To arrive at our final sample of $n = 249$ with valid data we inspected the screenshots and removed those who provided the incorrect screenshots or whose screenshots were illegible ($n = 37$).

Participants in the final sample were aged between 18 and 35 years old ($M = 24.07$, $SD = 5.01$), and 57.43% identified as Female, 40.56% as Male, and the remaining 2.01% identified with other gender descriptors. Given the limited age range, to facilitate group-wise comparisons, we dichotomized age into two categories, those younger than 25 years, and those 25 years or older. These two groups were labeled “GenZ” ($n = 146$, 58.63%) and “Millennials” ($n = 103$, 41.37%). In terms of education, the largest proportion had received a bachelor's degree (29.32%), while 32.13% had some college but no degree yet or only an associate's degree, 20.48% were high school graduates without any college experience, and 18.07% had a graduate or professional degree. In terms of racial makeup, our sample included more White ($n = 132$, 53.01%) than Asian ($n = 80$, 32.13%), Black/African American ($n = 18$, 7.23%), or Multiracial ($n = 11$, 4.42%) people. Participants who selected ‘other’ represented the smallest group in the sample ($n = 8$, 3.23%). Of the 249 people in the sample 41 (16.47%) identified as Hispanic.

Measures

The study was presented as an online survey and, during the first wave, we collected data on participants' demographics. For all four waves participants were asked to retrieve and donate screenshots from their iOS Screen Time report. The survey also collected data on various well-being related outcomes which do not form part of the present study.

To obtain behavioral data on the usage duration for applications used in the week preceding data collection, participants uploaded screenshots from their iOS Screen Time report. This feature, which is available on all iPhones running iOS v.12 or newer, passively tracks a variety of usage metrics. For this study participants were required to navigate to the “all activity” tab, select “week”, scroll to the “last week” section, and navigate to the view that shows the “most used” applications. This view displays a list of the most used applications in the previous seven days. For each application it provides the number of hours and minutes for which the application was used in the period. By default, the seven most used applications are displayed. Participants were asked to click “show more” and take a screenshot of the full list. To support participants in retrieving the correct data we provided video and textual instructions for how to navigate to the correct page and take and upload the screenshots.

To extract the data from these screenshots we developed a program that used *optical character recognition* to automatically convert the images to text. This involved extracting, firstly, the application name and, secondly, the number of usage hours and/or minutes from the screenshot. From the latter we calculated the total usage minutes for the preceding week for each application. Following an extensive cleaning process, we independently validated the accuracy of the conversion by checking the extracted data against the raw screenshot images. Any images that failed this check were manually extracted. These procedures resulted in a dataset, for each of the four waves, that included each participant’s most used applications in the week preceding data collection and the number of minutes for which these applications were used.

Findings

Which applications are most frequently adopted?

Across the four waves, 771 distinct applications were used by the participants. The number of users per application exhibited a power law relationship (bootstrap K-S = 0.02, $p = 0.33$), with just 10 applications accounting for 52% of all application-user combinations. Within each wave the number of distinct applications used was far less (wave 1: 401; wave 2: 367; wave 3: 378; wave 4: 369). Figure 1 summarizes the 15 most used applications across the full period and for each wave. *Safari* was consistently the most used application with close to 70% of the sample in each wave using it. While there were some minor changes, it is apparent that, at this aggregate level, application adoption was relatively consistent across the four waves.

Figure 1. Summary of the most used applications overall and per wave.

How much time do people spend using specific applications?

Table 1 provides key descriptive statistics for usage minutes for the top 15 applications with at least 15 distinct users, ranked by median usage time. With a median of 623 minutes per user per week across the four waves, *TikTok* was the application with the highest usage, despite being used by only 32% of the sample. In contrast, *Messages*, a far more popular application in terms of user numbers (78% of the sample used this application) was only the 13th most used application with a median of 149 minutes. Supporting these observations, for applications with at least 15 distinct users, the correlation between the number of distinct users and application usage duration was small ($r = 0.13$, $p < 0.001$, 95% CI [0.10, 0.15]). Extending Table 1, Supplementary Figure 1 summarizes the usage minutes for the most used applications in each wave. It is also apparent that *TikTok* is an outlier in terms of the proportion of participants' application usage time that it accounts for. All other applications with at least 15 users each account for less than 20% of a participants' application usage time. For these applications, the mean proportion accounted for by any one application is 11% ($SD = 5.2$). *TikTok*, however, accounts for 27.6% ($SD = 16.2$) of usage time for those who use the application, on average. These proportions were similar across waves (see Supplementary Table 1).

Table 1. Summary of the 15 most used applications in terms of weekly usage time across all four waves.

App	Users %	Usage time (mins)		% of participants' weekly usage
		Median	Mean (SD)	Mean (SD)
TikTok	32%	623.0	680.0 (441.0)	27.6 (16.2)
Instagram	72%	263.5	337.1 (259.0)	16.7 (12.2)
Among Us	12%	263.0	379.3 (403.7)	13.1 (9.8)
Safari	85%	262.0	403.8 (462.7)	17.2 (13.2)
YouTube	70%	260.0	475.4 (566.0)	18.7 (15.7)
Twitter	40%	259.5	329.8 (269.9)	15.4 (10.7)
Reddit	48%	240.0	355.6 (336.2)	16.8 (12.5)
Netflix	15%	204.0	439.7 (568.2)	12.7 (10.3)
Facebook	41%	193.0	273.9 (235.9)	14.5 (11.3)
Messenger	19%	174.0	406.7 (1042.3)	14.6 (16.2)
Twitch	9%	160.0	304.0 (363.3)	9.8 (8.4)
Discord	23%	159.5	234.8 (257.3)	11.1 (10.3)
Messages	78%	149.0	212.5 (188.4)	11.0 (8.4)
Chrome	22%	143.5	188.7 (181.5)	11.8 (11.9)
WhatsApp	10%	137.0	224.9 (288.2)	11.6 (10.7)

Figure 2 depicts an aggregated pairwise correlation network, produced using the ‘qgraph’ package (Epskamp et al., 2021), for weekly usage time for the 10 most popular applications. While usage time on some applications was strongly associated with time spent on other applications, on average, the within-person pairwise correlations among these 10 popular applications were small (mean $r = 0.06$). Exceptions include time on *Safari* which was positively associated with time on *Reddit* and *YouTube*, and time spent on *Facebook*, which was negatively associated with time on *TikTok* and *Snapchat*. Notably, most applications are pushed to the periphery of the graph instead of forming clusters (except for *Messages* and *TikTok*), illustrating that even among these 10 highly used applications, combinations are diverse rather than centralized. Interestingly, usage time on three popular image/video based platforms—*Instagram*, *TikTok*, and *Snapchat*—was only weakly related (all $r < 0.04$) in comparison with the associations observed among other applications.

Figure 2. A pairwise correlation network for application usage time for the 10 most popular applications. Blue lines represent positive associations, red lines negative ones, and the thickness and brightness of an edge indicates the magnitude of the association. The layout is based on the Fruchterman-Reingold algorithm and locates nodes with stronger and/or more connections closer together and the most central nodes in the center of the graph.

Do peoples' application choices differ according to demographic characteristics?

Table 2 summarizes the data on application choices by gender¹ and age across all four waves for the 10 most used applications in each group. For gender, while most applications exhibited similar usage levels, differences are apparent. Except for “GenZ” males, for whom *YouTube* had the highest adoption percentage, *Safari* was consistently the most used application. Compared to gender, the differences in application adoption were more pronounced for age group. *Facebook*, for instance, was used by 44% of those 25 or older, but only 21% of those younger than 25. In contrast, among this younger group, 35% used *TikTok*, while only 10% of older participants used *TikTok*.

¹For gender, given the small number of participants who used other gender descriptors ($n = 5$), we focused the analysis on those for which the sample contained sufficient data (male: $n = 101$; female: $n = 143$).

Table 2. Proportion of participants using an application by gender and age.

App	Male			Female			Age Group	
	GenZ	Millennial	Overall	GenZ	Millennial	Overall	GenZ	Millennial
Safari	58%	72%	64%	78%	76%	77%	69%	74%
YouTube	77%	42%	61%	55%	43%	50%	62%	42%
Instagram	56%	47%	52%	73%	55%	66%	66%	52%
Messages	53%	70%	60%	54%	64%	58%	54%	67%
Reddit	42%	51%	46%	27%	36%	31%	33%	42%
Snapchat	40%	8%	26%	44%		28%	42%	
Facebook	12%	36%	23%	28%	50%	37%	21%	44%
TikTok	25%		17%	42%	12%	30%	35%	10%
Twitter	44%	33%	39%	27%	15%	22%	35%	24%
Mail		24%	15%	11%	22%	15%		23%
Chrome	21%	23%	22%		13%		14%	17%
Discord	23%	16%	20%	15%		11%	18%	10%
Google.com		21%	13%	14%	16%	15%	11%	18%
Spotify	14%		11%	17%		11%	15%	

Figure 3 presents the level of adoption by indicating the proportion of participants within a given education category using each application. Except for those with only high school education, *Safari* was the most used application across education categories. For those with a

high school level of education, *Instagram* was the most adopted application with 71% of this group using this application.

Figure 4 depicts the 10 most adopted applications by gender in each of the four waves. Notably, *TikTok* only appeared in the 10 most used applications for males in wave 3, while it was in the top 10 of all four waves for females. Other notable differences include *Chrome* and *Discord* which, conversely, only made the top 10 for males, but not females.

Figure 3. Summary of application adoption by education level.

Figure 4. The 10 most adopted applications by gender in each of the four waves.

As shown in Figure 5, application usage across waves differed substantially by age category, with a number of applications only used by members of one group and not the other. *Snapchat*, *TikTok*, *Discord*, and *Spotify* only featured in the top 10 most adopted applications for those under 25 years old, while *Mail*, *Chrome*, and *google.com* only featured for those 25 and older. Additionally, a much larger proportion of younger participants used applications like *Instagram* and *YouTube* compared to their older counterparts.

Figure 5. The 10 most adopted applications by age category in each of the four waves.

Figure 6 depicts the 10 most used applications for each education category in each wave. Mirroring earlier trends, the top two applications were similar across education categories in terms of adoption percentage (*Safari*, *Messages*), but differences were apparent among applications ranked lower in the top 10. *Facebook*, for instance, exhibited higher levels of adoption among those with bachelors and graduate degrees compared to those with high school or some college. Similarly, *TikTok* did not feature in the 10 most adopted applications for those with graduate degrees, but it did for the three other categories.

Figure 6. The 10 most adopted applications by education category in each of the four waves.

Does application usage time differ according to demographic characteristics?

To further investigate demographic differences, we considered whether application usage time differs by gender, age group, or education level. Across all four waves, for applications with at least 10 users, females' per-application usage duration (median = 204 minutes, $M = 350.05$, $SD = 449.72$) was, on average, higher than males' usage (median = 152 minutes, $M = 252.31$, $SD = 312.51$), $t(5881.9) = 10.16$, $p < 0.001$, $d = 0.25$ 95% CI [0.20, 0.31]. This pattern was observed in all four waves (see Supplementary Figure 2). For both groups, *TikTok* was the application with the highest median usage duration overall and in all four waves. Notably, the difference in

correlations between adoption levels per application and usage duration per application for females ($r = 0.40, p = 0.03, 95\% \text{ CI } [0.05, 0.67]$) and males ($r = 0.31, p = 0.12, 95\% \text{ CI } [-0.08, 0.62]$) was not statistically significant, $z = 0.78, p = 0.43$.

Across all four waves, younger participants' usage duration per application (median = 188.5 minutes, $M = 326.94, SD = 446.44$) was, on average, higher than their Millennial counterparts (median = 173 minutes, $M = 286.66, SD = 319.33$), $t(5951) = 4.43, p < 0.001, d = 0.11, 95\% \text{ CI } [0.06, 0.16]$. For younger participants, across all four waves, *TikTok* was the most used application, with median usage surpassing the overall median usage. For older participants, usage patterns were far less consistent. *TikTok* was the most used application in wave 2 and wave 4, but did not make the top 10 in waves one and three. In these waves *Instagram* and *Safari* were the most used applications, respectively (see Supplementary Figure 3 for a complete summary). While the correlation between adoption levels per application and usage duration per application was not statistically significant for GenZ participants ($r = 0.25, p = 0.19, 95\% \text{ CI } [-0.13, 0.56]$), it was statistically significant for Millennial participants ($r = 0.49, p = 0.02, 95\% \text{ CI } [0.09, 0.76]$). The difference in correlations was also statistically significant ($z = -2.15, p = 0.03$).

Overall, per application usage was the highest among participants with only high school education (median = 213 minutes, $M = 357.54, SD = 533.52$) and was lower among those with some college (median = 176 minutes, $M = 311.97, SD = 389.88$), a bachelor's degree (median = 183 minutes, $M = 304.32, SD = 356.50$), or a graduate degree (median = 199 minutes, $M = 302.11, SD = 304.93$). Supplementary Figure 4 provides a complete account of these differences. Applications like *Instagram* and *YouTube* appear in the top five for all four categories, while *TikTok* has the highest median usage for those with high school diplomas, some college, or a bachelor's degree, but does not make the top 10 applications for those with a graduate degree.

How consistent are peoples' application choices over time?

To understand how stable participants' application usage was over time we calculated within-person "consistency scores." These scores were constructed by ranking the most-used applications for each person at each wave, calculating the absolute differences in the rank order of each application from wave to wave, then summing over all waves. The consistency score ranges from zero—signifying that the same five applications appeared in the same order at every wave—to 90, signifying that the five most-used applications were completely different from wave to wave. In addition to calculating summary statistics for the consistency score for the overall sample, we investigated whether there are differences across age category and/or gender.

Table 3. Within-person consistency of five most-used applications by gender and age category

	Mean (SD)	Median	Min-Max
Overall	31.4 (13.8)	30.0	6.0-72.0
Male			
GenZ	32.1 (12.5)	30.0	8.0-72.0
Millennial	31.3 (15.1)	24.0	6.0-68.0
Overall	31.8 (13.6)	28.0	6.0-72.0
Female			
GenZ	32.2 (14.6)	30.0	6.0-72.0
Millennial	29.6 (12.8)	28.0	8.0-71.0
Overall	31.1 (13.9)	30.0	6.0-72.0
Age Group			
GenZ	32.2 (13.7)	30.0	6.0-72.0
Millennial	30.3 (13.8)	28.0	6.0-71.0

Note. Consistency score is the sum of the absolute differences in the rank order of most-used applications from wave to wave at the within-person level. Lower scores represent more consistency, e.g., a score of 0 means that the same five applications appeared in the same top-five order across all four waves.

Summary statistics for the consistency score variable broken down by age category and/or gender are presented in Table 3 and the distributions are illustrated in Figure 7. Overall, every participant exhibited at least some fluctuation in their most-used applications, as evidenced by minimum values > 0 . An overall mean of 31.4 indicates that, on average, participants generally had a core group of most-used applications that fluctuated in their relative position in the top five, with some applications occasionally dropping out of and/or appearing in the top five. Looking across demographic categories, no (sub)group differed noticeably from the overall average consistency score. Similarly, average consistency scores for males versus females and

GenZ's versus Millennials were comparable. However, as evident in Figure 7, the median consistency score for male Millennials (24.0) was noticeably lower than other (sub)groups, suggesting that application use was generally more consistent for this group of participants.

Figure 7. Within-person consistency scores of five most-used applications by gender and age category. Lower scores signify more consistency over time. Dashed line signifies the overall median.

Does the amount of time that people use popular applications vary over time?

In addition to understanding the consistency of application rankings at the within-person level over time, we investigated whether the amount of time that people spend on certain popular applications varies within-persons across time. To do so, we calculated within-person means and standard deviations for the 15 most-used applications. As seen in Figure 8, applications with more variation in usage time from wave-to-wave were generally the applications with high(er) mean levels of usage time, demonstrating a strong linear pattern.

Figure 8. Average within-person mean and standard deviation (variation) for 15 most popular applications by usage minutes. Blue dashed line represents the line of best fit.

Discussion

Researchers tend to investigate smartphone usage as if it is a single uniform construct (c.f., Kaye et al., 2020). However, given that these devices afford users the ability to select their own mobile application repertoires, we need to focus on the emergent sets of applications that people use (Festic et al., 2021; Munger et al., 2021). Following Munger and colleagues' (2021) call for increased focus on the description of digital media usage, this study used longitudinal behavioral data collected via data donation to address many of the shortcomings in previous studies and investigate the types of applications that people commonly use, the amount of time they spend with these applications, the application combinations that they construct, and the consistency of these application combinations over time.

Overall, the results of this study suggest that peoples' mobile application repertoires are concentrated around a subset of popular applications that is relatively consistent over time. Just

30 applications account for over 70% of all recorded applications across participants. Outside of this group, and particularly outside of the 10 most used applications, a much smaller proportion of individuals adopt each application, and very little of their usage time (on average) is devoted to these applications. However, although usage is concentrated into a relatively small subset of applications, within this subset there is diversity between individuals and between applications. Taken together, these results suggest that quantifying smartphone usage with a single metric—total usage duration (i.e., ‘screen time’)—is unlikely to capture the full extent and diversity of media that users curate for themselves as part of their mobile application repertoires.

Our first research question concerned the applications most frequently adopted and, for these applications, the amount of time spent with them. The analysis revealed a set of applications that was consistently used by more than 50% of the sample (i.e., *Safari*, *Instagram*, *Messages*, and *YouTube*). These “core four” applications cover an array of important uses and gratifications characteristic of modern digital media—*Messages* for communication, *YouTube* for entertainment/information seeking, *Instagram* for self-expression, social interaction and entertainment, and *Safari* which, as a general-purpose web browser, enables access to a multitude of other constituent media and affordances—which may explain why they were so prevalent in our sample. Among the 10 most popular applications, while usage times were negatively correlated for some (i.e., *Facebook*, *TikTok*, *Snapchat*) and positively correlated for others (i.e., *Safari*, *YouTube*), in general time spent with these applications was only weakly related. The analysis also shows a relatively weak relationship between the level of adoption for an application and the amount of time people spend using it. This is particularly the case for *TikTok*, which had the highest usage time but was used by less than 1/3rd of the sample. *TikTok* was also an outlier in terms of the proportion of a participant’s usage time that it accounted for.

Generally, it accounted for close to 30% of individuals' application usage across waves, while all other popular applications accounted for between 11% and 20%. *TikTok*, therefore, exhibits greater 'stickiness' than the other applications described. Generally, this can be attributed to the content and the communication style inherent to the medium (Cuesta-Valiño et al., 2022).

Compared to other popular applications that generally emphasize interpersonal communication, *TikTok* primarily provides users with a personalized stream of algorithmically curated video content. Such algorithmic curation drives novelty and continued engagement (Scherr & Wang, 2021) and may, along with an interactive design that encourages active participation (Cuesta-Valiño et al., 2022), account for the substantial amount of time users spend with this application.

Our next research questions concerned the role of three demographic characteristics in mobile application usage. Our data indicate that, when it comes to the most adopted applications, a generally similar set of applications are used across these characteristics. Notably, however, considerable differences in the proportion of individuals adopting particular applications by age category were observed. For instance, the proportion of older people using *Facebook* is almost double the proportion of younger people who use the platform. Conversely, the proportion of younger participants using *TikTok* was almost three times as large compared to the proportion among older participants. Similar differences, albeit of a smaller magnitude, were observed for *Snapchat* and *Instagram*. That social media applications—particularly newer platforms like *TikTok*—were more prevalent among younger participants is consistent with recent self-reported data from a nationally representative sample of adults (Pew Research Center, 2021). On the other hand, email applications, *Google Maps*, *Twitter*, *Messages*, *Safari*, and *YouTube* were more popular among older participants. Considering these differences, it is apparent that, in this sample, younger participants preferred to use applications that afforded them access to and the

ability to share short-form video content, while older participants preferred more utilitarian (e.g., *Google Maps*) and general purpose (e.g., *Safari*) applications.

In contrast to the observed differences by age category, application choices were fairly similar between male and female participants. Despite this, in all four data collection waves, females' per-application weekly usage duration was on average higher than males' usage. Mirroring Kwon et al. (2021), notable distinctions emerge when looking at age differences *within* each gender category. For instance, nearly 75% of 'GenZ' females had *Instagram* as one of their most-used applications, while only 47% of 'Millennial' males did. Likewise, over 75% of 'GenZ' males had *YouTube* as one of their most-used applications, while only 43% of 'Millennial' females did. While these differences in application preferences may reflect, at least in part, the particularities of our convenience sample, the fact that *Instagram* and *YouTube* are popular among 'GenZ' is consistent with self-reported data from nationally representative samples (Pew Research Center, 2021). Our findings extend these by providing insight into how the popularity of these platforms varies by gender among 'GenZ' participants.

As the affordances and content available on these platforms are distinct, this has important implications for media effects research. Use of these platforms is likely to engender different media response states which may, in turn, lead to different effects (Valkenburg & Peter, 2013). Among other factors, differences in the types of media that constitute smartphone use may account for observed differences in effects reported in prior studies that only focus on overall measures of usage. For instance, certain studies have found that the association between media use and psychological distress is stronger for young females than other age-gender subgroups (e.g., Twenge et al., 2020). Although the reliability of this finding is questionable (Meier & Reinecke, 2020; Orben & Przybylski, 2020), if true, our findings suggest that

preferences in application repertoires across gender and age categories may play a role in observed differences in media effects. This further undergirds the importance of moving away from broad measures of screentime and toward more specific measures of media use.

Our final research question addressed the consistency of participants' application choices over time. Taken together, results of the within- and between-person analyses indicate that participants generally used the same set of applications over time, with some minor fluctuations in rank order. This finding extends prior research that found, over a shorter period (seven days), that smartphone users have unique and consistent behavioral patterns for use of 21 different applications (Shaw et al., 2022). Consistency in application usage may result from personal preferences, habituation, a lack of viable alternatives, market saturation, indifference to seeking alternatives. As many of the most frequently used applications afford users the ability to communicate with others or demonstrate a particular trait, identity, or characteristic, it is likely that a network effect exists—because a user's network of connections, friends, or followers is accessible through a particular application, not using that application would mean being cut-off from those connections. As in other online social networking contexts (Lewis et al., 2012), it is likely that an individual's application repertoire is, therefore, influenced by the choices of their peers. Further network-based analyses are required to assess this proposition

The relative stability in application use over time shown is noteworthy not only because the study spanned four months—a timeframe where one might expect some change in application choices over time—but also because the study period spanned the 2020 U.S. Presidential election (between waves 3 and 4), a momentous exogenous event or “shock” that could have hypothetically induced changes in participants' application choices (Van Aelst et al., 2021). These findings suggest that people engage in predictable application choices—at least

when it comes to their most-used applications—that are not easily disrupted by large-scale exogenous events such as an historic election or other ongoing changes during the study period. However, it will be important for future research to examine the consistency of application choices over longer time periods (i.e., years rather than months) and robustness to different types of exogenous shocks, such as change in employment status, residence, family composition, etc. In terms of usage time, not unexpectedly, we found that, on average, applications with more within-person variation in usage time from wave-to-wave were generally the applications with high(er) mean levels of usage time.

Limitations and Future Research

Although this study provides valuable insights into the application choices that drive people's mobile application repertoires and extends prior research that has primarily relied on self-report data to investigate adoption and usage, it is not without limitations.

A first set of limitations concern sampling and the degree to which the current findings are likely to generalize to the adult population at large. The data used in this study were collected from a relatively small convenience sample of iPhone users residing in the U.S. recruited through *Prolific*. Thus, future research is needed to determine the extent to which the findings generalize to smartphone users both within and outside of the U.S. While some similarities may exist, it is likely that, in addition to the applications highlighted in this study, users from other countries may adopt other applications that address local needs (e.g., *EskomsePush* in South Africa), are only available in specific regions (e.g., *Pinduoduo* in China), or have wider cross platform compatibility (e.g., *WhatsApp* compared to *iMessage*). To address these possibilities, there is a need for cross-country research with large, representative samples to leverage data donation approaches and investigate cross-country differences in mobile application repertoires.

A second set of limitations concern the data collected. While recent research has highlighted the discrepancy between self-reported and behavioral data on smartphone usage (Parry et al., 2021), digital trace data are not without their own limitations and biases (Jürgens et al., 2019). Most notably, as we relied on data donations in the form of screenshots in this study, we were unable to independently verify the accuracy of the data. There is a need for research to validate the accuracy of these logging features and determine the degree to which they are calibrated to actual usage behavior. In addition to uncertain accuracy, the form of the data limited the amount of insight we could gain into application usage. While the current study focused on specific application usage patterns—an improvement over approaches that simply focus on the overall duration of usage—there remains a need for further insight into the purposes for which applications are used and the content interacted with. In particular, many applications (e.g., web browsers or meta-applications like *WeChat*), like smartphones, are themselves meta-media that afford users access to a range of uses, content, and interactions. To fully characterize smartphone usage, research will need to focus on the motivations and purposes driving usage of narrow-purpose applications as well as these more general broad-purpose applications.

Conclusion

This study demonstrates the potential of longitudinal data donation of iOS Screen Time reports for the investigation of mobile application usage. The study has generated important descriptive findings about the mobile application repertoires that individuals develop and shown how application usage centers around a small core of applications that various (sub)groups augment in particular ways. The study highlights the consistency of application choices and draws attention to demographic differences in these choices. Across these findings, the central contribution of the study is the description of the diverse set of media that constitute smartphone

use. When studying effects associated with smartphone use, it is no longer tenable to aggregate use of all these distinct applications into a single high-level measure (Kaye et al., 2020; Parry et al., 2022). Smartphone usage is a diverse and heterogeneous behavior that should be treated as such. In future research that builds on this study, researchers should strive to gain a deeper insight into, firstly, the content and purposes characterizing application usage and, secondly, the motivations driving application usage through qualitative research methods.

References

- Boczkowski, P. J., Matassi, M., & Mitchelstein, E. (2018). How Young Users Deal With Multiple Platforms: The Role of Meaning-Making in Social Media Repertoires. *Journal of Computer-Mediated Communication*, 23(5), 245–259.
<https://doi.org/10.1093/jcmc/zmy012>
- Boeschoten, L., Ausloos, J., Moeller, J., Araujo, T., & Oberski, D. L. (2020). Digital trace data collection through data donation. *ArXiv:2011.09851 [Cs, Stat]*.
<http://arxiv.org/abs/2011.09851>
- Büchi, M. (2020). *Digital Well-Being Theory and Research* [Preprint]. Open Science Framework. <https://doi.org/10.31219/osf.io/k3e2j>
- Cuesta-Valiño, P., Gutiérrez-Rodríguez, P., & Durán-Álamo, P. (2022). Why Do People Return to Video Platforms? Millennials and Centennials on TikTok. *Media and Communication*, 10(1). <https://doi.org/10.17645/mac.v10i1.4737>
- Dickson, K., Richardson, M., Kwan, I., Macdowall, W., Burchett, H., Stansfield, C., Brunton, G., Sutcliffe, K., & Thomas, J. (2018). *Screen-based activities and children and young people's mental health and psychosocial wellbeing: A systematic map of reviews*. Department of Health Reviews Facility.
<http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3748>
- Epskamp, S., Costantini, G., Haslbeck, J., Isvoranu, A., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2021). *qgraph: Graph Plotting Methods, Psychometric Data Visualization and Graphical Model Estimation* (1.9).
<https://CRAN.R-project.org/package=qgraph>
- Ferguson, D. A. (1992). Channel repertoire in the presence of remote control devices, VCRs and

- cable television. *Journal of Broadcasting & Electronic Media*, 36(1), 83–91.
<https://doi.org/10.1080/08838159209364156>
- Festic, N., Büchi, M., & Latzer, M. (2021). How Long and What For? Tracking a Nationally Representative Sample to Quantify Internet Use. *Journal of Quantitative Description: Digital Media*, 1. <https://doi.org/10.51685/jqd.2021.018>
- Freelon, D. (2014). On the Interpretation of Digital Trace Data in Communication and Social Computing Research. *Journal of Broadcasting & Electronic Media*, 58(1), 59–75.
<https://doi.org/10.1080/08838151.2013.875018>
- Griffioen, N., Scholten, H., Lichtwarck-Aschoff, A., van Rooij, M., & Granic, I. (2021). Everyone does it—differently: A window into emerging adults’ smartphone use. *Humanities and Social Sciences Communications*, 8(1), 177.
<https://doi.org/10.1057/s41599-021-00863-1>
- Griffioen, N., van Rooij, M., Lichtwarck-Aschoff, A., & Granic, I. (2020). Toward improved methods in social media research. *Technology, Mind, and Behavior*, 1(1).
<https://doi.org/10.1037/tmb0000005>
- Haight, M., Quan-Haase, A., & Corbett, B. A. (2014). Revisiting the digital divide in Canada: The impact of demographic factors on access to the internet, level of online activity, and social networking site usage. *Information, Communication & Society*, 17(4), 503–519.
<https://doi.org/10.1080/1369118X.2014.891633>
- Hasebrink, U., & Popp, J. (2006). Media repertoires as a result of selective media use. A conceptual approach to the analysis of patterns of exposure. *Communications*, 31(3).
<https://doi.org/10.1515/COMMUN.2006.023>
- Heeter, C. (1985). Program Selection With Abundance of Choice. *Human Communication*

- Research*, 12(1), 126–152. <https://doi.org/10.1111/j.1468-2958.1985.tb00070.x>
- Horvát, E.-Á., & Hargittai, E. (2021). Birds of a Feather Flock Together Online: Digital Inequality in Social Media Repertoires. *Social Media + Society*, 7(4), 205630512110528. <https://doi.org/10.1177/20563051211052897>
- Humphreys, L. (2018). *Smartphones as Metamedia: A Framework for Identifying the Niches Structuring Smartphone Use*. 17.
- Jensen, K. (2016). Metamedium. In *The international encyclopedia of communication theory and philosophy* (pp. 1–4). John Wiley & Sons.
- Jones-Jang, S. M., Heo, Y.-J., McKeever, R., Kim, J.-H., Moscovitz, L., & Moscovitz, D. (2020). Good News! Communication Findings May be Underestimated: Comparing Effect Sizes with Self-Reported and Logged Smartphone Use Data. *Journal of Computer-Mediated Communication*, 25(5), 346–363. <https://doi.org/10.1093/jcmc/zmaa009>
- Jürgens, P., Stark, B., & Magin, M. (2019). Two Half-Truths Make a Whole? On Bias in Self-Reports and Tracking Data. *Social Science Computer Review*. <https://doi.org/10.1177/0894439319831643>
- Kaye, L., Orben, A., A. Ellis, D., C. Hunter, S., & Houghton, S. (2020). The Conceptual and Methodological Mayhem of “Screen Time.” *International Journal of Environmental Research and Public Health*, 17(10), 3661. <https://doi.org/10.3390/ijerph17103661>
- Kim, S. J. (2016). A repertoire approach to cross-platform media use behavior. *New Media & Society*, 18(3), 353–372. <https://doi.org/10.1177/1461444814543162>
- Kwon, S. E., Kim, Y. T., Suh, H., & Lee, H. (2021). Identifying the mobile application repertoire based on weighted formal concept analysis. *Expert Systems with Applications*, 173, 114678. <https://doi.org/10.1016/j.eswa.2021.114678>

- Lewis, K., Gonzalez, M., & Kaufman, J. (2012). Social selection and peer influence in an online social network. *Proceedings of the National Academy of Sciences*, 109(1), 68–72.
<https://doi.org/10.1073/pnas.1109739109>
- Matassi, M., Mitchelstein, E., & Boczkowski, P. (2022). Social media repertoires: Social structure and platform use. *The Information Society*, 1–14.
<https://doi.org/10.1080/01972243.2022.2028208>
- Meier, A., Domahidi, E., & Günther, E. (2020). Computer-Mediated Communication and Mental Health: A Computational Scoping Review of an Interdisciplinary Field. In S. J. Yates & R. E. Rice (Eds.), *The Oxford Handbook of Digital Technology and Society* (pp. 78–110). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190932596.013.4>
- Meier, A., & Reinecke, L. (2020). Computer-mediated communication, social media, and mental health: A conceptual and empirical meta-review. *Communication Research*, 1–28.
<https://doi.org/10.31234/osf.io/573ph>
- Munger, K., Guess, A. M., & Hargittai, E. (2021). Quantitative Description of Digital Media: A Modest Proposal to Disrupt Academic Publishing. *Journal of Quantitative Description: Digital Media*, 1. <https://doi.org/10.51685/jqd.2021.000>
- Ohme, J., Araujo, T., de Vreese, C. H., & Piotrowski, J. T. (2021). Mobile data donations: Assessing self-report accuracy and sample biases with the iOS Screen Time function. *Mobile Media & Communication*, 9(2), 293–313.
<https://doi.org/10.1177/2050157920959106>
- Orben, A. (2020). Teenagers, screens and social media: A narrative review of reviews and key studies. *Social Psychiatry and Psychiatric Epidemiology*, 55(4), 407–414.
<https://doi.org/10.1007/s00127-019-01825-4>

- Orben, A., & Przybylski, A. K. (2020). Reply to: Underestimating digital media harm. *Nature Human Behaviour*, 4, 349–351. <https://doi.org/10.1038/s41562-020-0840-y>
- Parry, D., Davidson, B. I., Sewall, C. J. R., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour*, 1–13. <https://doi.org/10.1038/s41562-021-01117-5>
- Parry, D., Fisher, J. T., Mieczkowski, H., Sewall, C. J. R., & Davidson, B. I. (2022). Social Media and Well-being: A Methodological Perspective. *Current Opinion in Psychology*, 45. <https://doi.org/10.1016/j.copsyc.2021.11.005>
- Pew Research Center. (2021). *A majority of Americans say they use YouTube and Facebook, while use of Instagram, Snapchat and TikTok is especially common among adults under 30*. 19.
- Scherr, S., & Wang, K. (2021). Explaining the success of social media with gratification niches: Motivations behind daytime, nighttime, and active use of TikTok in China. *Computers in Human Behavior*, 124, 106893. <https://doi.org/10.1016/j.chb.2021.106893>
- Schumacher, S., & Kent, N. (2020, April 2). *8 charts on internet use around the world as countries grapple with COVID-19*. Pew Research Center. <https://www.pewresearch.org/fact-tank/2020/04/02/8-charts-on-internet-use-around-the-world-as-countries-grapple-with-covid-19/>
- Sewall, C. J. R., Goldstein, T. R., & Rosen, D. (2021). Objectively measured digital technology use during the COVID-19 pandemic: Impact on depression, anxiety, and suicidal ideation among young adults. *Journal of Affective Disorders*, 288, 145–147. <https://doi.org/10.1016/j.jad.2021.04.008>

- Sewall, C. J. R., & Parry, D. (2021). The Role of Depression in the Discrepancy Between Estimated and Actual Smartphone Use: A Cubic Response Surface Analysis. *Technology, Mind, and Behavior*, 9.
- Shaw, H., Taylor, P. J., Ellis, D. A., & Conchie, S. M. (2022). Behavioral Consistency in the Digital Age. *Psychological Science*, 7.
- Taneja, H., Webster, J. G., Malthouse, E. C., & Ksiazek, T. B. (2012). Media consumption across platforms: Identifying user-defined repertoires. *New Media & Society*, 14(6), 951–968. <https://doi.org/10.1177/1461444811436146>
- Twenge, J. M., Haidt, J., Joiner, T. E., & Campbell, W. K. (2020). Underestimating digital media harm. *Nature Human Behaviour*, 4(April), 346–349. <https://doi.org/10.1038/s41562-020-0839-4>
- Valkenburg, P. M., & Peter, J. (2013). The Differential Susceptibility to Media Effects Model: Differential Susceptibility to Media Effects Model. *Journal of Communication*, 63(2), 221–243. <https://doi.org/10.1111/jcom.12024>
- Van Aelst, P., Toth, F., Castro, L., Štětka, V., Vreese, C. de, Aalberg, T., Cardenal, A. S., Corbu, N., Esser, F., Hopmann, D. N., Koc-Michalska, K., Matthes, J., Schemer, C., Sheafer, T., Splendore, S., Stanyer, J., Stępińska, A., Strömbäck, J., & Theocharis, Y. (2021). Does a Crisis Change News Habits? A Comparative Study of the Effects of COVID-19 on News Media Use in 17 European Countries. *Digital Journalism*, 9(9), 1208–1238. <https://doi.org/10.1080/21670811.2021.1943481>
- Walsh, L. C., Okabe-Miyamoto, K., Regan, A., Twenge, J., & Lyubomirsky, S. (2021). *The Association Between Well-Being and Objectively Measured Versus Self-Reported Smartphone Time* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/zcwmv>

Wolf, C., & Schnauber, A. (2015). News Consumption in the Mobile Era: The role of mobile devices and traditional journalism's content within the user's information repertoire.

Digital Journalism, 3(5), 759–776. <https://doi.org/10.1080/21670811.2014.942497>

Yuan, E. (2011). NEWS CONSUMPTION ACROSS MULTIPLE MEDIA PLATFORMS: A repertoire approach. *Information, Communication & Society*, 14(7), 998–1016.

<https://doi.org/10.1080/1369118X.2010.549235>