Temporality of Political Content Exposure via Smartphone Screens:
A Computational Description of Rapidity and Idiosyncrasy

A dissertation submitted to the Department of Communication

and the committee of graduate studies of Stanford University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Daniel Muise

May 2022

Abstract

This dissertation provides a computational description of the rapid and idiosyncratic nature of exposure to political content in the smartphone environment, in order to critically address implicit conceptual and methodological assumptions common in computational political communication research. The data used are hyper-rich longitudinal screen-recordings collected from 115 Americans' smartphones for up to two weeks each in late 2019, amounting to over four years of natural screen usage. Drawing on theoretical temporal foundations of empirical research, I discuss the importance of timescale explication in political communication research and the temporally fragmented nature of smartphone usage. I then discuss how political content fits into the smartphone environment, in terms of format, sequencing, psychological processing, and especially information retention. Tying together these considerations, I demonstrate measurement validity concerns that arise from (1) aggregating unique encounters with political content as commensurable units, (2) aggregating durations of exposure to political content as commensurable units, and (3) utilizing singular formats or sources (including the *news* format) as proxies for measuring political content exposure. I also find connections between political content exposure durations and idiosyncratic content sequencing as afforded by the smartphone's flexibility. Throughout, I illustrate and quantify high variance in core measures both across and within subjects, and I show that the risk of ecological fallacy in this domain is similar interindividually and intraindividually. I conclude with an assessment of current research practices, suggestions for the field, reflection on limitations, and directions for future research.

Table of Contents

Section 1. Introduction

Smartphone users can access political content at any time, for any duration, from any location, in essentially any format (Molyneux, 2018; Reeves et al., 2019; Shim, You, Lee, & Go, 2015). The incredible affordances of the smartphone, and flexible media technology more generally, have enabled a radical shift in the range of experiences available to news audiences, starkly dissimilar to the preceding century. Whereas the broadcast era had an incredibly small menu of news sources, the present era has a virtually infinite source array. Whereas the cable era had fixed programming timeslots and channels, the present era allows for fragmentation of news exposure across the day. The possibilities available to political communication researchers have greatly expanded as well. Modern digital technology produces massive amounts of digital trace data, enabling measurement of natural media interactions far more precisely than self-reports and laboratory settings allow (Freelon, 2014; Jungherr, 2018; Lazer, 2020).

Given this new media environment and its corollary data resource, certain topics in political communication have garnered special attention across popular and academic discourse. Chief among these is mass polarization, along with its hypothesized catalytic systems based in flexible media technology: echo chambers and filter bubbles (Guess, Nyhan, Lyons, & Reifler, 2018). These topics come into direct contact with the lived experience of many Americans and are built on long-held academic concerns of the media failing to properly inform the voting public (Pariser, 2011; Sunstein 2004, 2017; Terren & Borge, 2021). Yet, empirical research on filter bubbles, echo chambers, and technology-fueled polarization using digital trace data has yielded surprisingly mixed results regarding the intensity or even existence of these phenomena (Bruns, 2017; Dubois & Blank, 2018; Guess et al., 2018). The preponderance of large-scale observational computational studies have revealed that the typical American consumes very little news via their smartphone or social media (Allen, Howland, Mobius, Rothschild, & Watts, 2020), and moreover is not in any sort of echo chamber or filter bubble based on their exposure and sharing behavior (Bakshy, Messing, & Adamic, 2015; Barberá, 2014; Dubois & Blank, 2018; Eady, Nagler, Guess, Zilinksy, & Tucker, 2019; Flaxman, Goel, & Rao, 2016; Guess et al., 2018; Guess, 2021; Morales, Borondo, Losada, & Benito, 2021; Muise, Hosseinmardi, Howland, Mobius, Rothschild, & Watts, 2022; Peterson, Goel, & Iyengar, 2018). What are we to make of this gap between theoretical expectations and popular understanding on the one hand, and our still-unsettled empirical evidence on the other, despite seemingly perfect observational data? I put forward that conventional techniques of using digital trace data to measure exposure to political content are invalid due to the continued application of latent outmoded measurement assumptions tied to prior media eras but not appropriate for the study of flexible media technology usage. Specifically, there are three closely-linked outmoded assumptions latent in the bulk of recent research.

First is the aggregation of counts of unique encounters with political content. The conventional approach of treating unique instances of exposure as commensurable units of analysis is not appropriate for the study of modern media and is rather a holdover from prior eras of mass communication, wherein programming occurred in uniform blocks of time (Lazer, 2020; Tewksbury, Weaver, & Maddex, 2001). The temporal flexibility afforded by smartphones allows users to access political content for any duration with any level of attentiveness. Skimming, scrolling, and switching-back to non-political content can occur at will and requires only an instant (Matthes, Nanz, Stubenvall, & Heiss, 2020). Hence, upon a unique encounter with political content, media users may stay tuned for hours or they may not actually engage with any amount depth or attentiveness, representing very different user experiences (Lazer, 2020; van Damme, Martens, Van Leuven, Vanden Abeele, & De Marez, 2020).

Second is the aggregation of exposure durations into total exposure times. Measurement strategies which aggregate political content exposure without regard for the duration of individual encounters risk equating entirely different psychological processes per the experience of the smartphone user (Lemke, 2000). The processing style of political content encounters depends on the speed at which processing occurs (Lodge & Taber, 2005; Stoker, Hay, & Barr., 2020), as well as the user's goals upon information encounter (Kaye & Johnson, 2004), and the user's willingness to expend time on elaboration (Van Damme et al., 2020). Thorough processing that leads to actual knowledge gain or opinion change requires an expenditure of time and cognitive resources by the media user (de Zúñiga, Borah, Goyanes, 2021). Even outside of a distracting media environment, rational citizens are not motivated to expend cognitive resources toward processing political content, beyond personal interest or intrigue (Downs, 1957). More generally, political content exposure elicits different processes at different speeds and durations (Matthes et al., 2020). Studies of repeated exposure that do not differentiate between durations are ignoring such differences.

Third is the measurement of news exposure as a proxy for exposure to political content. With the modern decoupling of political content from its older formats, *politics* is a content-based descriptor of media genre, not a format-based one. In contrast, the term *news* primarily describes a source or formatting style that was once the hegemon delivery mechanism for political content, based on conventions in the business of journalism. News formats no longer comprise the modal instance of political content exposure, as I will show, and so measurement strategies erroneously affixed to news articles or formal sources risk measuring not political content but something else entirely. Research that does not identify this change in the field cannot adequately describe exposure to political content in the present media environment (Matthes et al. 2020; Ryfe, 2019).

In all three cases — aggregation of unique exposures, aggregation of exposure time, and conflation of formats with political content — vast interindividual and intraindividual variation further complicates all measurement. Flexible temporality is a media affordance adopted idiosyncratically by each user at each occasion of their lives. Unchecked assumptions of homogeneity across research subjects, and of stability within subjects over time, are ill-advised in the study of flexible media usage, and especially in the study of smartphone usage (Reeves, Robinson, & Ram, 2020). Moreover, the role of sequencing in political content engagement is unclear, but recent evidence of self-regulatory behavior on the smartphone suggests that it cannot be ignored (Cho, Reeves, Ram, Robinson, & Yang, 2022). The methodological importance of temporal agency in the smartphone environment is further compounded by the high likelihood of researchers inadvertently sharing singular subjects through popular convenience sampling tools.

Scientific progress in the field of political communication requires a clear understanding of audiences' experiences. Missing from the literature is a richly informed ground-truth understanding of the smartphone user's experience of political content, in all its natural complexity. In this dissertation, I fill this gap by conducting a rigorous description of the rapidity and idiosyncrasy of political content exposure on the smartphone. In line with an emerging goal in social science, and specifically digital media studies, I provide rich quantitative description as a research outcome, with an aim to theoretically ground the ongoing study of complex evolving systems through rigorous empirical baseline inquiry (Jebb, Parrigon, & Woo, 2017; Munger, Guess, & Hargittai, 2021). To accomplish this goal, I use uniquely hyper-rich observational data to measure the durations, frequencies, and interindividual and intraindividual variations of encounters with political content on the smartphone, as guided by crucial gaps and unfounded assumptions latent in the field.

My data was collected under the Screenomics paradigm, which is a passive and rapid smartphone screenshot logging technique (Reeves et al, 2021). This data provides multiple advantages for the present task. Screenomics data cuts across format distinctions (such as news/non-news) by capturing all screen activity agnostically with respect to software applications. It also provides access to natural behavior outside of laboratory environments through passive collection in the field. Lastly, it provides a seconds-level temporal granularity, necessary for studying quick exposures (Reeves, et al., 2020). The dataset I use was collected from 115 American research subjects in 2019, with each subject contributing between six and fourteen days of screenshot log data, summing to over four years of human experience.

The rest of this dissertation is structured as follows. First, in Section 2 I explain how time, timescale, and temporal measurement theoretically fit into empirical research. In Section 3, I examine how timescale is currently incorporated into political communication research, specifically in studies of mass polarization. In Section 4, I reexamine the *building blocks* of our field, news and political content, which have become obfuscated in the smartphone environment. In Section 5, I reexamine the rapid and idiosyncratic environment itself, and the experience of political content exposure within it. In Section 6, I introduce four research questions that address crucial unknowns about how political content is experienced on the smartphone. In Section 7, I describe the unique screenshot data used for this study and its value for the goals of my dissertation. In Section 8, I describe the variables that I extract from my raw data in order to conduct analysis. In Section 9, I present my analysis and results according to the research questions posed in Section 6. In Section 10, I organize and discuss the implications of my findings, catalog limitations of my dissertation, and offer directions for future research. A supplementary appendix is provided with additional information, analyses, and visualizations.

Section 2. Temporal measurement, timescales, and intraindividual variation

In this section, I provide theoretical background for my dissertation by explaining the role of timescale and its explication in empirical social science research.

The smartphone's flexibility hinges on the personalized fragmentation of experience across time into small bits (Reeves et al., 2019). To understand the implications of fragmentation for the psychological processing of content (and its measurement), a proper explication of timescale is required. The timescale of a process is the period over which a process or phenomenon occurs, and thus is measurable.

Temporal foundations

Timescales are best understood with reference to the fundamental components of empirical social science. There are three fundamental axes of empirical data, which Cattel (1952) refers to as factors. These are *variables*, *persons*, and *occasions*, as illustrated in Figure 1 Panel A. The first of these factors, variables, regards the topic of research. The second of these factors, persons, has garnered ample attention in the field of political communication with the advent of digital trace data. Larger cross-sectional sample sizes imply greater statistical power for determining effect sizes (Cohen, 1992) and greater likelihood of producing cross-sectionally generalizable results (Brennan, 1992; Yarkoni, 2022). The third factor, occasions, highlights how all phenomena are processes unfolding through time, exhibiting new states at new moments. Whereas variables and persons are generally discrete units, an occasion is an arbitrarily bounded discretization of a continuous timeline.

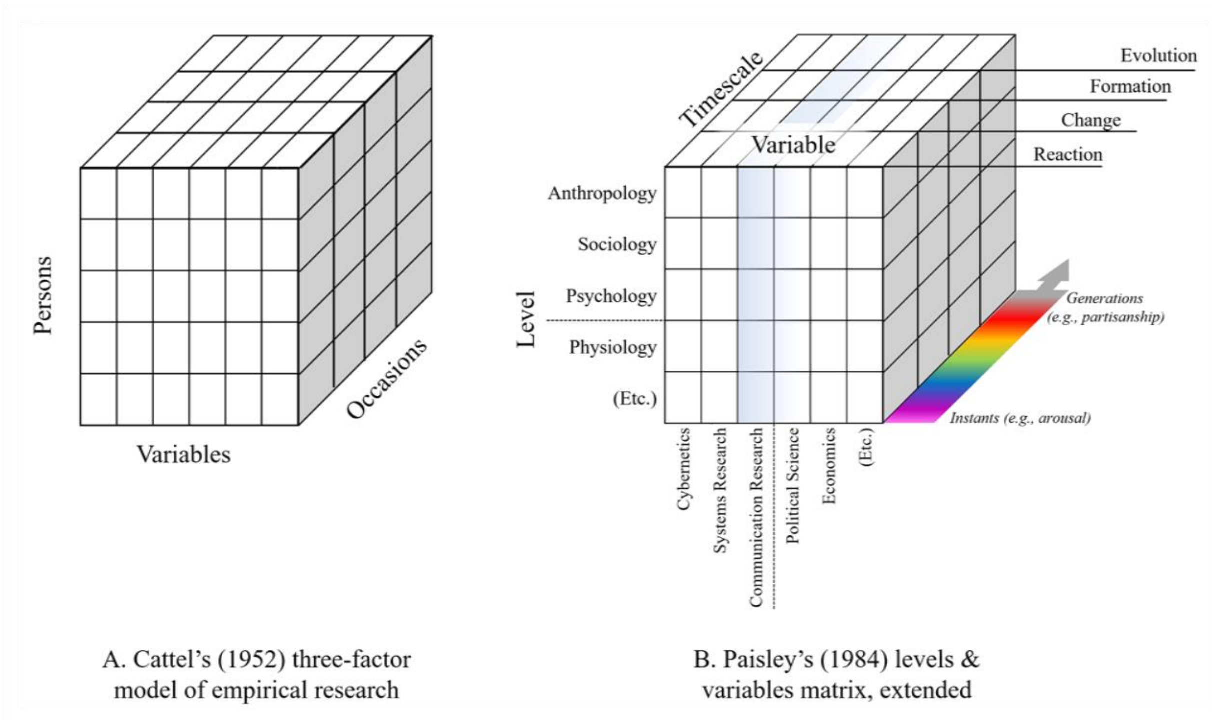


Figure 1A. An illustration of Cattel's (1952) three factors of empirical research. Cross-sectional research is represented by the front surface of this cube, wherein variation across occasions is held constant or otherwise unattended to. Idiographic research is represented by the top surface of the cube, wherein inter-personal variation is not attended to. Only panel data allows for the incorporation of all three factors.

Figure 1B. Building on Paisley's (1984) illustration of "Communication within the Social Science" levels-variables matrix (with clear similarities to Paisley's earlier work), I add a third dimension demonstrating timescales of communication concepts as an independent axis from level and variable. Timescales range from instants to generations, and potentially longer spans. Coloration is modeled after wavelengths in the visible spectrum to highlight the periodic nature of phenomena in time.

The mapping from occasions to timescales is analogous to the mapping between persons and cross-sectional levels of analysis. To help illustrate this point, I expand on Paisley's (1984) levels-variables matrix in Figure 1 Panel B, which bears a striking resemblance to Cattel's (1952) foundational empirical model. Paisley's levels-variables matrix (the frontward face of the cube in Figure 1 Panel B) categorizes social science into disciplines. Along the bottom axis are several variable-analogous labels, construed in the broadest sense as entire fields of study. Each of these variables is manifest and measurable at multiple cross-sectional levels, shown along the vertical axis. Level of analysis determines the cross-sectional size of the sample that must be analyzed for a whole phenomenon to be observed. For example, while a psychological phenomenon can be observed in a single individual, an anthropological phenomenon can only be observed in a critically large mass of humans. Note that communication and political science (and thus their border, political communication) span multiple levels of cross-sectional analysis; phenomena of interest to political communication scholars can occur at the level of an individual or a society.

To Paisley's matrix I add an orthogonal third axis of *timescale* over which phenomena may occur and be measured. The timescale degrees from reaction to evolution were chosen arbitrarily to demonstrate stopping points along a potentially infinite temporal spectrum. Phenomena occurring at the timescale of a generation cannot be measured by data spanning only one day; phenomena occurring over the timescale of an instant cannot be measured by data whose occasions are sampled once every hour. In his original matrix, Paisley implicitly treated timescale as a *component* of a cross-sectional analysis level, given its logical correlation with the more tangible concept of cross-sectional scale (Bronfenbrenner & Morris, 2006). However, in the present media environment, timescale and cross-sectional scale are no longer tightly coupled. In the smartphone era, communication processes at massive cross-sectional levels can occur in an instant, e.g., global reaction to a single tweet. Likewise, communication processes occurring at the intra-human physiological cross-sectional scale can last for a lifetime, e.g., through habitual interaction with the smartphone screen (Lazer et al, 2009; Miller, 2012).

Timescale correspondence between process and measurement

Like cross-sectional levels of analysis, timescales are each composite of one another (Lemke, 2000; Molenaar, 1985; Ram & Diehl, 2015). A shift in an individual's partisan-biased opinion year-on-year is the composite of faster phenomena occurring within a year – e.g., a news diet that fluctuates month-on-month (Muise et al., 2022a). Month-on-month changes to the news diet are themselves the composite of daily news exposure experiences, comprised of single television news show durations, and so on down the milliseconds. News exposure measurements captured in the course of a single month would be a poor tool studying processes whose period lasts an entire year or more, especially considering the likelihood of month-on-month fluctuations. This measurement mismatch is within scope of Lemke's (2000) Adiabatic Principle: processes occurring at a given timescale are more influential on processes at timescales of similar wavelength than on processes occurring at more dissimilar wavelengths (i.e., much longer or much short timescales).

In the case of a researcher using a one-month study to investigate a year-long process, the researcher must assume that their findings resulting from the study of news exposure *within a month* would generalize across time to the *span of a year*. This assumption is more commonly explicated in the use of cross-sectional data used to generalize to persons outside the sample. Symmetrically, a *year-on-year* measurement of an individual's changing partisan identity may be appropriate for describing the individual at the yearly level, but such a measurement strategy cannot adequately describe values occurring *between* intervals. In practice, however, change in news exposure (or similar variables) between two measurement occasions is conventionally assumed to be linear, which is not necessarily the case (e.g., Muise et al., 2022a).

At finer timescales, incremental changes in time measurement are proportionally more extreme. Sending an annual survey five seconds late may not alter estimates at a yearly interval, but for processes lasting five seconds, an additional five seconds would double the time available for processes to unfold. Conversely, a process lasting for ten seconds cannot be measured in a five second span. More fundamentally, a process lasting ten seconds cannot *occur* in a five second span. This deceptively simple point raises concerns for political communication research, as I discuss in the next section.

Variation along the temporal axis

Figure 1 also illuminates the very important possibility of a temporal ecological fallacy. The ecological fallacy is the erroneous assumption that group-level estimates apply to singular units within a group (Piantadosi, Byar, & Green, 1988). Group-level estimates, primarily the statistical average, summarize variant values of a random variable. Means and other summaries have the potential to not represent the experience of any unit in the group. Similarly, intraindividual values of a random variable can fluctuate across occasions, meaning that, e.g., the average monthly behavior of an individual may not represent behavior in any month of the individual's life. Researchers analyzing variables cross-sectionally generally might not consider such fluctuations, assuming that cross-sectional values are not importantly dependent on time or are randomly distributed through time or across subjects, but this need not be the case. Muise et al., (2022a), examining the scale and evolution of echo chambers in the American news audience, uncovered a consequential case of temporal ecological fallacy in the literature. Using granular panel data, they found that even though the share of Americans who consume partisan-skewed new diets stayed very stable each month between 2016 and 2019, the intraindividual news diets of apparent echo chamber members were actually changing month-to-month in their partisan composition, meaning that a newly reshuffled share of the American population had partisan-skewed news diets at each cross-sectional aggregation. In that context, the *apparent* intraindividual stability of biased news consumption was at odds with the genuine intraindividual variation, greatly altering the substantive interpretation of findings.

Just as intraindividual variance is often unacknowledged relative to interindividual variance, intraindividual *covariance* between variables is often overshadowed or proxied by interindividual covariance. Similarity between covariances within the same bivariate relationship is referred to as ergodicity (Molenaar, 2004; Molenaar & Campbell, 2009; Salvatore & Valsiner, 2010). Hamaker (2012) provides an example of how easily ergodicity assumptions can be broken, using typing speed and typographical error rate. Measuring across individuals, typing speed is negatively correlated with typographical errors; measuring within individuals, typing speed is positively correlated with typographical errors. The relationship between typing speed and typographical error rate is *non-ergodic*, meaning that intraindividual experience cannot be learned or inferred through cross-sectional analysis. The potential for non-ergodicity is underexplored in political communication research, as with other social sciences (Fisher at al., 2018), requiring data with high temporal resolution.

Even to the extent that audience-level aggregate measurements are interpreted appropriately (i.e., not applied intraindividually), reporting on cross-sectional outcomes alone masks crucial intra-group extremes and intraindividual dynamics which may be far more meaningful or consequential (Molenaar, 2008). The well-known diversity of inter- and intraindividual political interest (Boulianne, 2011; Prior, 2010), access to political content (Bimber, 2001), and political participation or efficacy (Kenski & Stroud, 2006; Leighley &

Nagler, 1992) make clear that broadness of cross-sectional aggregation is negatively associated with informational utility of analysis beyond broad summarization goals.

In the next section, I tie this theoretical backdrop to research in the field of political communication.

Section 3. Temporality in political communication research

In this section, I examine timescale and temporality explication in political communication research, drawing on studies of mass polarization using digital trace data.

The prior the section outlined core concepts of timescale and temporal measurement in empirical social science research. In this section, I narrow this discussion by focusing on temporality as it arises in political communication.

Conceptual treatment of temporality in political communication

Many political communication concepts are poorly explicated in the literature, with various definitions and operationalizations, and numerous blurred boundaries (McLeod et al., 2009; Scheufele, 2000). This lack of explication percolates to explication of both timescale and cross-sectional levels. Where conceptual boundaries are blurred across cross-sectional levels of analysis, researchers commonly aggregate individual-level phenomena to explain or predict audience-level shifts and trends (Price, Ritchie, & Eulau, 1991). In contrast, the aggregation of processes across time is less explicitly documented, with assumptions of timescale independence leading to naïve extrapolations of results from fast processes to longer-term processes, and vice-versa. The assumption that measurements of quick exposures, behaviors, and reactions of individuals sum or scale linearly across timescales is unfounded and potentially misleading.

Consider for example that attempts to measure partisan echo chambers have often referenced the concept of cognitive dissonance (via the selective exposure it motivates) to explain ongoing societal shifts in news exposure behavior (Metzger, Hartsell, & Flanagin, 2020; Winter, Metzger, & Flanagin, 2016). Not only does this multi-tiered system span cross-sectional levels, but also timescales: cognitive dissonance is a near-instantaneous sensation (Elliot & Devine, 1994; Festinger, 1957), selective exposure concerns repeated action based on stable preference (Bryant & Zillmann, 1984), while an echo chamber is definitionally a longer-term process (Nguyen, 2020), and still slower is the timescale of polarization, which echo chamber research is most often motivated by.

When referring to polarization, I am referring to *mass polarization* literature, regarding the increasing separation of Americans along the generally linear spectrum of political ideology. This is an intensely popular topic in 21st century research, following from the apparent bifurcation of partisan opinion and growing intensity of partisan identity in the U.S. population (Baldassarri & Bearman, 2007; Lelkes, 2016). The study of mass polarization, its origins, and its consequences are of great interest to scholars of political communication, who are quick to seek associations between political media and changes in public opinion, with special emphasis on the role of the Internet (Druckman & Levendusky, 2019; Muise et al., 2022a; Prior, 2013). Given the vast array of approaches to researching this massive topic, the timescales of mass polarization and its constituent processes consist of a broad array of timescales with ambiguous interdependence that is rarely explicated.

At its greatest extent, mass polarization is an intergenerational process, perhaps the longest process currently studied empirically in political communication. Acceleration of co-partisan marriage preference and filial partisanship transmission have been theorized in the past and confirmed by a measurement strategy spanning fifty years of data (Davies, 1965; Iyengar, Konitzer, & Tedin, 2018). Still, the apparent expansion of mass polarization in the modern era has inspired investigation into countless communication processes at shorter timescales, each assumed to be partially culpable for increases in audience-level cross-partisan hostility (Barberá, 2020).[1] Primarily, the dual concepts of filter bubbles (Pariser, 2011) and echo chambers (Garret, 2009a) have become central to political communication research, under the assumption that both are systems that catalyze polarization within the public (Flaxman, Goel, & Rao, 2016). Neither concept comes with clear delineation of timescale, or temporal relationship to the overall polarization trend. Filter bubbles describe the restrictive impact that personalized algorithms have on the diversity of political viewpoints encountered by individuals (Bechmann & Nielbo, 2018; Pariser, 2011; Thorson, Cotter, Medeiros, & Pak, 2021). A filter bubble requires sufficient time for an algorithm to *incubate*, or train on a user's preferences, and then sufficient time to alter the views of its members (Haim, Graefe, & Brosius, 2018). Echo chambers don't imply any such algorithmic incubation period. Instead, they have been treated as static traps where co-partisans are stuck in prolonged communication, at the relative exclusion of contrasting viewpoints (Colleoni, Rozza, & Arividsson, 2014; Dubois & Blank, 2018; Morales et al., 2021).

Research into these dual concepts has been done at both the individual level and audience level and has studied their causes, their effects, and more often, their existence. Out of all

[1] Not discussed here are several polarization-related research programs outside the bounds of communication research, such as partisan-ideological sorting (Davis & Dunaway, 2016), social identity patterns (West & Iyengar, 2020), and elite polarization (Banda & Cluverius, 2018).

hypothesized causes for echo chambers and filter bubbles, selective exposure has most often been pointed to as a key behavioral culprit (Stroud, 2010). Selective exposure is a moment-by-moment choice to expose oneself only to congenial information, or the circumstantial inevitability of this phenomenon due to homophily in society. Hence, selective exposure refers to repeated occurrences of fast decisions occurring over multiple time points across long spans, with deleterious effects on individual-level political knowledge (Freedman & Sears, 1965; Sears & Freedman, 1967). Selective exposure has been studied meticulously throughout the history of political communication, usually at the level of singular choices and instantaneous cognitive motivations and reactions (Garret, 2009, Garrett, 2013; Hart, Albarracín, Eagly, Brechan, Lindberg, & Merrill, 2009).[2] In the context of polarization, the process of selective exposure has been assumed to aggregate linearly across time, at the audience level over weeks or months (e.g., Bakshy, Messing, & Adamic, 2015; Gentzkow & Shapiro, 2011) and the individual level where behavior is examined in seconds or minutes (Albarracín & Mitchell, 2004; Arceneaux, Johnson, & Murphy, 2012). This leaves open several questions around temporal aggregation: is any instance of political content exposure equally important as, and independent from, the rest? How can a brief moment of exposure be meaningfully aggregated with an hour-long exposure (Matthes et al., 2020; Van Damme et al., 2020)?

Consequences of timescale obfuscation in polarization research

[2] Selective exposure has also often been studied in longer-term settings, mainly in in other communication fields, e.g., selective exposure to movies (Weaver, 2011), television programming (Bryant & Zillmann, 1984), and radio (Best, Chmieliewski, & Kreuger, 2005).

Investigation into the mere *existence* of online echo chambers has produced a variety of descriptive analyses using impressively massive cross-sectional audiences. These studies often gather data from long durations but then aggregate across time (Barberá, Jost, Nagler, Tucker, & Bonneau, 2015; Bessi, 2016; Boutyline & Willer, 2017; Dubois & Blank, 2018; Eady et al., 2019; Garrett, 2009, 2013; Gentzkow & Shapiro, 2011; Guess, 2015; Guess, 2021; Kim, 2011; Liao & Fu, 2014; Peterson et al., 2019). Others have measured echo chambers according to proliferation processes over months (Del Vicario et al., 2016), patterns of moment-level individual choices over longer spans (Flaxman, et al., 2016), or the evolution of audience-level news exposure over years (Muise et al., 2022a; Yang, Majo-Vazquez, Nielsen, & González-Bailón, 2020). While it is tempting to approach echo chambers as though they are static, and therefore able to be studied cross-sectionally, the theoretically essential component of an echo chamber or filter bubble is its longevity as experienced by any one individual (Muise et al., 2022a). Exposure to diverse viewpoints requires time enough to be exposed sequentially in an unbiased manner. Cross sectional analyses of whole audiences frequently risk capturing nonsensical 'visits' into echo chambers, wherein an individual's news diet is only temporarily biased over the timespan of analysis. This temporal gap renders many studies of echo chambers somewhat fruitless in their attempt to explain the longer trend of polarization (Muise et al., 2022a).

The primary empirical concern is one of measurement validity: what is it that we measure when we measure an instance of *exposure* or *engagement* such as a click, view, like, share, scroll or retweet (Lazer, 2020; Matthes et al., 2020)? For instance, consider how digital trace measurements of exposure and social engagement are over-employed by the seventeen relevant studies I have referenced thus far.[3,4] All seventeen studies explicitly or implicitly drew conclusions about *engagement* with political content online, but apart from Muise et al. (2022a), no study has employed minimum time thresholds when measuring exposure to political content, thus assuming an unrealistic commensurability between momentary glimpses and long-lasting engagement — i.e., different spans into which different psychological processes may fit. Only three of seventeen studies incorporated exposure *duration* information at all (Allen et al., 2020; Guess, 2021; Muise et al., 2022a). The remainder assumed all unique URL visits carry equal weight (Flaxman et al., 2016; Gentzkow & Shapiro, 2011; Guess et al., 2018), or that unique exposures and/or sharing behaviors are an appropriate proxy for the actual (temporal) experience of media users (Bakshy et al, 2015; Barberá, 2014; Barberá et al., 2015; Bechmann & Nielbo, 2018; Colleoni et al., 2014; Conover et al., 2011; Eady et al, 2019; Morales et al., 2021; Munger et al., 2017; Nithyanand, Schaffner, & Gill, 2017; Peterson et al., 2021). Several of these studies had the data available to measure durations of exposure but operationally treated each unique political content encounter as being identically received and processed (Lazer, 2020).

One potential impact of this approach is the likely *under*estimation of the extent of partisan segregation in the political audience, via the scale of filter bubbles or echo chambers. To the extent that political content is presumed to *impact* users in some way, no measurement strategy benefits from assuming that 51 momentary exposures to counterpartisan information outweigh 50 hour-long exposures to pro-partisan information. More foundationally, 50 unique momentary exposures to political content should not be weighed 50-fold more impactful than a

[3] (Allen et al., 2020; Bakshy et al, 2015; Barbera, 2014; Barbera et al., 2015; Bechmann & Nielbo, 2018; Colleoni et al., 2014; Conover et al., 2011; Eady et al, 2019; Flaxman, Goel, & Rao, 2016; Gentzkow & Shapiro, 2011; Guess, 2021; Guess et al., 2018; Morales et al., 2021; Munger et al., 2017; Nithyanand et al., 2017; Muise et al., 2022; Peterson et al., 2021)
[4] Note that smartphone-specific studies using digital trace data are still comparatively uncommon in this domain.

single thirty-minute exposure, without clear explication of how such a method impacts outcome measurements. With current research, we do not know if these arbitrary examples are edge cases or the norm. In the following two sections, I build out these theoretical ramifications through a targeted discussion of the smartphone environment and what its dynamics mean for the study of political content exposure. The first step is to clarify the evolving format of political content and how its decoupling from the *news* format has obfuscated recent research efforts.

Section 4. Distinguishing News from Political Content

In this section, I discuss how the outdated use of the term "news" within the fragmented modern environment further complicates measurement strategies and validity.

As basic building blocks of political communication, the terms *news* and *political content* demand explication as a first step toward understanding the temporality of political content exposure in the present media environment. A century ago, *news* referred to information from beyond the senses, distributed mainly through print and radio. The broadcast television era borrowed this functional definition, and the *news desk* and *news bulletin* visual and temporal formats were held from prior eras all the way through to modern cable news (Hamilton, 2005). Through this entire trajectory, news formats were intimately tied to the central *topic* of news: political content (Eveland & Seo, 2009). Hence, beyond local happenings (including weather and crime), to read the news, listen to the news, or to watch the news was to consume political content.

The tradition of coupling political content with the news format is no longer realistic. A basic example of this decoupling would feel familiar to any smartphone user. Suppose a user is scrolling through the frontpage of a social media application when she lay eyes on a so-called *hard news* article written by a seasoned journalist. This user keeps scrolling. A second later, she sees an earnest post about a conspiracy theory from a distant relative. Then, at the same time, an involuntary Amber Alert message might appear at the top of her screen, while an email notification shows a monthly crime report from the neighborhood watch. Before she puts her phone down, her brother calls to tell her about his political views, her sister shares a link about an event happening in the area, and then her friend texts her a link to a celebrity-gossip article. Which of these experiences are news? In five minutes, she has taken in a great deal of information about the world, especially the political world, but only one piece of content came in a standard news format, i.e. an article by a professional journalist.

In the research literature, news has been variously defined, both implicitly and explicitly, by authors from disciplines ranging from computer science to communication theory. Papers that define news by its *function* focus on its informative value or functional utility. In this view, news and its demand are biological remnants of our human interest in surveillance. Researchers using this definition imply that modern news is founded on newsworthiness, that is, actual utility for consumers' well-being — and that the news market happens to be flooded with stories painted with the biological signals of surveillance value: sensationalism, salaciousness, and so on. Thus, news is whatever portrays itself as useful surveillance about the world outside (Bennett, 2003; Reeves, 1989; Shah, Watts, Domke, Fan, & Fibison, 1999; Shoemaker, 2006; White, 1950).

Some research applies an alternative *industrial* definition, wherein content is *news* if it was produced by a proper news-producer, e.g., the New York Times or Fox News (Bakshy et al., 2015; Baum, 2003; Baum & Groeling, 2008; Chakraborty et al., 2016; Dilliplane, Goldman, & Mutz, 2013; Flaxman et al., 2016; Messing & Westwood, 2014; Shoemaker, 2006; Yeykelis et al., 2014). Yet other research articles combine function and industry (Galtung & Ruge, 1965; Harcup & O'Neill, 2001; Iyengar et al., 1984; Molotch & Lester 1974; 1975; Mondak, 1995; Shoemaker, 2006). Some research also makes use of topical cues when subsetting content. This is more often true of quantitative or web-based papers which define news by focusing on articles A) produced by top media outlets and b) including semantic markers (Althaus, 2002; Althaus & Tewksbury, 2002, Flaxman et al, 2016; Guess et al, 2018; Munger et al, 2017; Nithyanand et al., 2017; Prior, 2003). One final definition, audience-based, is implied through large-scale quantitative studies and computer science. The best example of an audience-driven definition is from Kwak et al. (2010), which sought to determine if Twitter was a social network or a news outlet. Under this concept, news is information that chiefly gets consumed without reciprocal information being sent to the source. Similar implications can be found in Friggeri, Adamic, Eckles, & Cheng (2014) and Prior (2003).

I take this academic divergence in news definitions and operationalizations to imply a separation between user experience and vestigial heuristic concepts of what news is. This alone motivates a closer look at the exact experience of modern-day political media consumers. In this paper, I proceed by stepping back from the ambiguous and spacious term news, and instead focus on *political content*. The boundaries of political content are strictly topical. Though the precise content covered under the term *political content* may feel intuitive, this intuitive sense can be difficult to pin down. Much as with the term *news*, an individual's interpretation of the term's boundaries may vary according to particular worldview or experience. What sets political content apart, however, and suitable for the smartphone environment, is that it may be broken down into immutable criteria based on a singular content definition conveyed in any format. In Section 8, I discuss operationalization of this definition in the context of my data, where I focus on ideologies, individuals, institutions, and events directly related to governance and social justice. In the next section, I employ this refocusing on the *topic* political content to examine its presence and dynamics in the complex and rapid smartphone environment.

Section 5. Fast-Thinking and Rapidity of Political Content Exposure on the Smartphone

In this section, I discuss how the smartphone environment enables interactions with political content at high speeds with unclear consequences on users, relative to prior media regimes structured for considerably longer-lasting instances of exposure.

Smartphones and other flexible media technologies have up-ended a variety of traditional media formats. As just described, news and political content are among the seemingly stable conventional concepts whose boundaries have become obscured in recent years. At the root of this change and complexity is the newly ubiquitous affordance to customize the timing of media experience. Digitization and the flexibility of the smartphone jointly allow for the consolidation of virtually any task or topic onto a single screen, each fragmented across time to suit moment-level changes in a user's goals, interests, or emotional states (Brinberg et al., 202l; Oulasvirta, 2005). The result for each user at each moment is a stream of rapidly sequenced content and behavior that is utterly idiosyncratic, impossible for a human to fully recall, and essentially irreplicable in human history (Reeves et al., 2021).

What does this environment mean for the study of political communication? Returning to the building blocks of our field: the inherent value of exposure to political content is the presumed impact that it has on the exposed. In considering any sort of impact on the exposed, we must consider that an impact is the result of a process. Content takes *time* to process, and different processes occur at different timescales. On a smartphone, the difference between five and fifteen seconds differentiates an inattentive scroll-by on a social media news feed from a sizable pause in rapid scrolling behavior (Brinberg et al., 2020). In aggregating any measure of content exposure in the fast-paced flexible media environment, researchers must assume some degree of homogeneity in the processes unfolding at the various timescales of the various exposure durations to which users are exposed. So then, what is known about processes of information exposure, consumption, engagement at timescales afforded by flexible media technology?

Rapid processes in political communication

The role of fast thinking and the rapid sequencing of topics has been familiar to communication scholars for decades under different terminology and in different domains (e.g., Miller & Campbell, 1959). The *priming effect* is the influence that some priming stimulus *A* has on a receiver's reaction to some subsequent message *B* (Berkowitz & Rogers, 1986; Jo & Rogers, 1994). Priming effects are in a sense presumed inevitable, in that they occur without conscious thought by the person experiencing them. For a receiver to subconsciously associate the prime with the message, the entire one-off process must occur within the short span of human attention (Domke, Shah, & Wackman, 1998; Ghuman, Bar, Dobbins, & Schnyer, 2008). Thus, the effect is restricted to timescales of minutes or less, and research into the priming effect has been well-suited to the laboratory (Hermans, De Houwer, & Eelen, 2001). In political communication, applications of priming most often concern impressions of candidates and news stories based on the chronological ordering of news broadcasts (Roskos-Ewoldsen, Klinger, & Roskos-Ewoldsen, 2007). The influence of the priming effect at longer timescales is not well-understood, and hence, in political communication literature, it has been assumed to linearly compound in long-term natural usage, as in the case of media violence (Jo & Berkowitz, 1994).

The priming effect was conceptualized in an era of linear media, where the sequencing of content was largely out of a media consumer's control. In contrast to linear media (such as television commercials and feature films), flexible media allow for tight personal control over when 'scene cuts' occur (Lang, Zhou, Schwartz, Bolls, & Potter, 2000). Yeykelis et al. (2014) measured moment-by-moment skin capacitance of laptop users and found that a spike in arousal is consistently just *before* users switched tasks, in contrast to just *after* a scene cut in linear media. The implication is that task-switching behavior (and by extension, compulsive smartphone checking) satisfies a rapidly arising and physiologically measurable impulse within the device user (Oulasvirta, Tamminen, Roto, & Kuorelahti, 2012). The smartphone attends to a common desire for rapid content changes and thus short durations of exposure to any single topic, including political content.

Thus, mental conditions necessary for elaborative thought that are not commonly associated with smartphones and social media (Cacciatore et al., 2018; de Zúñiga et al, 2020; van Erkel and Van Aelst, 2020). One named condition is the "prepared mind", or a mental state open for engagement with complex topics (Kouinos et al., 2006; Rubin, Burkell, Quan-Haase, 2011). A prepared mind requires a degree of mindfulness and a lack of urgent distractions and compulsions (McBirnie, 2008). Indeed, information processing of any kind requires the allocation and expenditure of finite cognitive resources, something to which humans are evolutionarily averse absent sufficient motivation (Lang, 2009). The smartphone's array of enticing low-cognitive-load behavior choices does little to prepare a mind for meaningful engagement with political content (e.g., Leung, 2020), where distractions may be an incoming notification, a post on a scrolling feed just peeking past the bottom of the screen, or the physiologically arousing awareness that new experience is just one tap away (Yeykelis, Cummings, & Reeves, 2014).

Fast & slow thinking

Daniel Kahneman separates mental processes into two speed-based modes: "fast thinking" and "slow thinking" (2011). These modes are not binary but are relative descriptors on a spectrum of cognitive depth. Fast thinking refers to shallow and intuitive comprehension built upon mental heuristics and impressions without deliberation, e.g., seeing a scowling face and sensing the emotion it conveys, or encountering a fan of a rival politician and forming an instant judgement about how you may get along. By its nature, fast thinking occurs largely automatically. In contrast, slow thinking lasts at least multiple seconds and involves active engagement with stimuli, e.g., determining the reasons behind a face's scowl or evaluating redeemable qualities of your political rival.

As fast thinking is less cognitively demanding, humans are universally inclined to save energy by applying fast thinking where possible, and shift to slow thinking only when necessary or interesting (Fiske & Russel, 2010). This effort-based distinction between fast and slow thinking is directly analogous to the peripheral and central routes (respectively) for persuasion in the Elaboration Likelihood Model (ELM) (Petty & Cacioppo, 1986). The ELM posits that an individual who is not likely to be convinced through elaborative channels (i.e., slow thinking) may instead be convinced through less cognitively demanding "peripheral" communication. For instance, while a politically-engaged voter may eagerly think through their vote choice based on an evaluation of policy positions, a less-engaged voter may select a candidate at the voting booth based on name recognition. Stoker, Hay, & Barr (2020) specifically examine how fast and slow thinking influence voters' impressions of politics as a whole. Using a focus group design where group members were sequentially induced to think in either mode, they found that fast thinking resulted in near-universal negativity shared with high confidence, but that slower thinking consistently resulted in net-neutral attitudes and group agreement that politics is a necessary and complex process.

Matthes et al. (2020) theorize how fast and slow processing jointly determine political content engagement outcomes during a social media usage session. According to their model of *Political Incidental News Exposure* (PINE), an individual who encounters political content incidentally in their content feed (e.g., as an article, a meme, a comment thread) may either deem the political content relevant or irrelevant to their own interest. Relevant political content switches the user's information goal from their initial impression (fast thinking) toward intentionally engaging with the political content, i.e., slow thinking and elaboration. If the user deems the political content irrelevant, then the political content is scrolled-past, leaving only a small and short-lived impression on the user according to fast-thinking processes. The PINE model was developed to understand political content engagement on social media but applies to the entire smartphone experience. Incidental exposure is incredibly common in smartphone usage, perhaps even as the primary source of political encounters on the platform (Dimmick, Feaster, & Hoplamazian, 2011; Van Damme, 2020; Weeks, Lane, Kim, Lee, Kwak, 2017). If this is the case, the informational utility of political engagement on the smartphone is not clear, as the smartphone environment is designed in opposition to elaborative thinking. As also described in the PINE model, an individual who is intentionally seeking political content can be distracted by encounters with other political topics or non-political topics (Matthes et al, 2020). In either case, the same relevancy rule applies: if a user determines that the incidental encounter is relevant, the user will abandon their original information goal; if not, the user's engagement with political content will have been irreversibly influenced by fast mental processes associated with the distraction (Matthes et al., 2020).

Information retention & sequencing

On smartphones, the psychological and temporal micro-context into which political content is embedded is inextricable from how it is experienced, with exposure duration being at the center. It is beyond the scope of this dissertation to directly measurement how the micro-context of political content exposures impacts specific outcomes. Rather, I focus on one universally relevant measure of how information exposure impacts the exposed: retention. Message elaboration cannot occur if the content on which to elaborate does not exist in memory (Singh, Rothschild, & Churchill, 1988). Table 1, adapted from Goldstein, McAfee, & Suri (2011), shows the sharp decline in recall and recognition of information content as exposure duration approaches zero. These recall and recognition rates are based on surveys taken by subjects just minutes after exposure to online advertisements. Of these, the value most pertinent to the smartphone environment is at the bottom left of the table. Advertisements viewed for five seconds which were second in a sequence of advertisements on a single screen had a combined recall and recognition rate of just 12%. Conversely, 97% of subjects exposed to an advertisement for four minutes (Table 1, top right) recognized the content of the advertisement they were exposed to. Crucially, the gradation between these extremes is not linearly additive in time.

Seconds of exposure to advertisement:	**5**	**10**	**25**	**40**	**60**	**120**	**180**	**240**
Content Recognition (%)	25	28	43	43	75	85	89	97
Content Recall (%)	12	16	24	29	48	63	62	66
Recognition & Recall, ad shown 1st in sequence (%)	31	33	43	46				
Recognition & Recall, ad shown 2nd in sequence (%)	12	14	28	32				

Table 1. Recognition and recall of the content of on-screen advertisements according to duration of exposure; data is adapted from Goldstein, McAfee, & Suri (2011). Recall and recognition rates are estimated based on surveys taken by subjects a few minutes after exposure. In the first two rows, results from exposures durations greater than 40 seconds are based on exogenously imposed durations; exposure durations less than 60 seconds are based on endogenously determine viewing durations. In the latter two rows, exposure durations are exogenously determined, and advertisements were shown in timed sequences, with recognition and recall averaged together. R&R rate for 5-second ad exposures second in a sequence was just 12%.

Other advertising research provides additional insight into how shortening viewing durations at the seconds-level timescale versus longer ads (Dhote & Kumar, 2019) reduces the impact on the viewer, in terms of persuasion ability (Dunst, 1993) and brand appraisal or likeability (Rogers 1995; Patzer, 1991). In this dissertation, I focus primarily on political content exposure segments lasting on the order of a few seconds. Not only are they the most prevalent segments in my data (described in my analysis), but as Table 1 indicates, empirical research that aggregates such segments with longer-duration segments may be mixing two different types of experience.

While the smartphone environment enables short durations and fast task-switching behavior, underlying motivations are somewhat more complex, but can be understood in part as a tool for emotional regulation or rapid mood management (Cho, et al., 2022). In that study, smartphone sequencing behavior is regarded as an ongoing process of self-regulation. Applying the circumplex model of arousal and valence (Russell, 1980) to Screenomics data highly similar to the sample used in this dissertation, the authors find that smartphone users balance arousal states sequentially according to the counter-regulation principle (Rothermund, 2011). As a matter of automatic affective processing, a usage session replete with content that is highly arousing (e.g., an explicative-laden text conversation) is significantly more likely to be followed by a session composed of low-arousal content (e.g., an instruction manual) and vice versa. They find a similar self-regulatory behavior regarding the *informational* content of screenshots. As the authors detail, their finding corroborates similar studies conducted in the laboratory setting: humans attend more readily to emotional stimuli that contrast with preceding emotional states and stimuli (Derryberry, 1993; Schwager & Rothermund, 2013). Similarly, there is some evidence that humans attend more readily to stimuli contrasting with overarching emotional *norms* in society (Hsu, Niiya, Thelwall, Ko, Knutson, & Tsai, 2022). In the American context, high arousal and positive valence form the dominant emotional norm (Tsai, Knutson, & Fung, 2006). Media users have an attentional bias toward counter-normative, i.e., negatively-valenced content (Hsu et al., 2022). Political content tends to have a more negative valence than non-political topics (Quattrone & Tversky, 1988; Soroka, 2006) and is often temporally embedded in comparatively positively valenced content. This generally suggests a system in which political content is attended to for longer durations following segments of positively valenced content, but there is as yet no evidence in this domain.

Altogether, the fast-paced smartphone environment, with its task-switching affordances and infinite variety across persons, occasions, and topics, is a highly complex medium for exposure to political content, not well-suited for the unchecked application of conventional aggregation or detection methods. In the following section, I outline a research strategy for studying the temporality, rapidity, and idiosyncrasy of political content exposure this environment.

Section 6. Research Questions

In this section, I lay out a research strategy for developing ground-truth understanding of political content exposure in the smartphone environment, motivated by the theoretical and empirical discussions in the preceding sections, and following a framework of rigorous quantitative description.

The complexity and intractability of the smartphone environment has rendered the actual experience of its users opaque to political communication researchers. The lack of dedicated description respecting the rapid and idiosyncratic nature of smartphone usage represents a skipped step in the research program of our field. What is needed is a clear and rigorous description of smartphone usage experience at the individual level, such that future research in this domain can find solid footing in clearly documented behavior (Jebb et al, 2017; Munger et al. 2021). My dissertation is poised to fill this gap by asking and answering the research questions listed throughout this section, drawing on the theoretical discussions of the preceding sections.

First, to establish the baseline prevalence of encounters with political content on the smartphone, I first aim to measure the frequency of these encounters, in terms of unique instances and overall occurrence.

> *RQ1.1 What is the frequency with which political content appears on the smartphone screen?*

From this, I can then identify how patterns of political content encounters vary across individuals. These variations are often entirely overlooked in aggregate measurements of audiences but are central to assessments of the idiosyncratic behavior enabled by smartphones. Should political content exposure be highly varied across individuals, extrapolation from group-level estimates to the individual-level would be an instance the ecological fallacy.

RQ1.2 How does political content exposure frequency vary across individuals?

Intraindividual variation is central to the uniqueness and temporality concerns of measurement in the smartphone environment. Using standardized methods, intraindividual variation can be directly compared to interindividual variation, thus illuminating how well individual-level aggregate values capture any particular moment in time.

RQ1.3 How does political content exposure frequency vary within individuals?

Next, I study the durations of political content. Prior literature suggests that durations of political content exposure can be very short-lived relative to conventional long-form political media 's thirty-minutes blocks. I aim to provide direct evidence as to the degree of this phenomenon, which could determine the manner in which political content is typically processed psychologically when encountered.

RQ2.1 What is the distribution of durations of political content segments on the smartphone?

I then turn attention to how duration distributions vary across individuals. Strong variation suggests that different users may gain differently from political content encounters, even given a common *frequency* of unique encounters.

RQ2.2 How do political content segment durations vary across individuals?

Similarly to RQ 1.3, I then examine intraindividual variation in duration distributions. Conventional wisdom suggests that some individuals with greater interest in political content may spend longer durations attending to political content, but this is unclear. Examining intraindividual variation in durations also allows for investigation of ergodicity in bivariate relationships regarding political content consumption, i.e., how duration distributions are related to overall political content exposure inter- and intraindividually.

RQ2.3 How do political content segment durations vary within individuals?

Next, I study how these frequencies and durations vary across applications. Smartphone users have countless sources for political content, including interactive communication, social media, news applications, and so on. The field lacks direct evidence of how political content exposure is distributed across these sources in natural usage, interindividually and intraindividually.

RQ3.1 How do political content segment frequencies vary across application categories?

Similarly, there is a lack of direct evidence of how political content segment durations are distributed across sources on the smartphone. If some sources consistently retain users' attention for much longer time spans than do others, then this would suggest an important bias in the quality of political content exposure, possibly with strong interindividual variation.

RQ3.2 How do political content segment durations vary across application categories?

Lastly, I investigate the political content exposure durations according to valence, arousal, and wordcount, as measured within segments and in terms of sequence effects. Drawing on Cho et al. (2022), I particularly study whether self-regulation processes are discernable in political content exposure duration variation. I use the valence-arousal circumplex model (Russel, 1980) to evaluate how political content segments' sentiment and wordcount, and *preceding* sentiment and wordcount, correlate with the longevity of political content on the smartphone screen. Should any of these features be related to the segment duration, this would suggest that processes related to segment duration (such as retention) are related to the rapid and inherently idiosyncratic sequencing of content on the smartphone.

RQ4 How do segment durations vary according to valence, arousal, and wordcount preceding and during segments?

In the following section, I describe the unique data used in my dissertation analysis and motivate its use for these particular research questions.

Section 7. Screenomics Data / Sample Description

In this section, I describe my sample and explain its utility for my dissertation goals.

My dissertation goals require high temporal resolution and clarity of screen activity that is entirely agnostic to the source and content of information appearing on the smartphone screen. To that end, I use data collected according to the Screenomics paradigm, which maintains that a proper understanding of the experience of fragmented media segments can only be ascertained through direct observation of natural behavior (Reeves, Robinson, & Ram, 2020). The conventional application of this paradigm is to remotely and passively collect screenshots from the smartphone screen every five seconds (and immediately upon screen activation) while the smartphone screen is in use. This process is accomplished through the Stanford Screenomics Lab's custom-built Screenomics data-collection application for Android phones. Screenshots are captured as PNG files, and these PNG files are minimally compressed in JPG files before being uploaded through secure encryption to a single cloud server. All data collection and analysis was conducted with approval of the Stanford University Institutional Review Board (protocol 38485). Raw screenshot files were stored exclusively on approved cloud storage or associated virtual machines and accessed only in accordance with up-to-date human subjects training and certification. Along with these image files, the data collection application provides a separate log of all instances of updates to the application in the screen’s foreground during regular usage. Text is extracted from each frame using task-specific optical character recognition, or OCR (Chiatti et al., 2017).

The 115 research subjects used in this dissertation analysis were recruited through the Qualtrics panel aggregation service in 2019, using a non-intersectional screening-based demographic quota system to approximate representation of the United States population, including geographic dispersal across the country. These individuals filled out a survey which included a set of screening questions confirming that they owned a compatible phone (Android operating system, exclusively used, and not shared) and a declaration of their interest in participation for two weeks of smartphone screen-recording in exchange for $30 incentive payment.[5] Survey respondents who agreed to this arrangement were forwarded onto a Stanford-hosted website on which they could provide consent according to a consent form approved by the Stanford University Institutional Review Board,[6] and then follow instructions for installing the Screenomics data collection application on their own smartphones. Upon confirmation that an individual’s smartphone was uploading screenshots and application log data following installation, a survey respondent was considered a subject actively enrolled for an intended two-week period. Table 2 shows a summary of subjects’ contributions to the raw data used in this dissertation.

[5] Recruitment survey available in Appendix K.
[6] Consent for available in Appendix L.

	Total	Mean	Minimum	Median	Maximum
Subjects	115				
Screenshots	4,907,091	42,670	1,069	37,819	178,094
Screentime (hours)	6,815	59.3	1.5	52.5	247.4
Subject Days	1,463	13	6	14	14
Daily Screentime (hours/day)		4.6	0.1	4.0	17.7

Table 2. Description of the sample. Rows indicate sample features. Screenshots are the basic unit of data, though not always the unit of analysis in this dissertation. Screentime, shown in hours, is the approximate total screen activity of all subjects estimated through screenshot count. Subject-days is the number of days of smartphone usage included in the subsample. The sample is trimmed such that all subjects contribute a minimum of 6 consecutive days wherein the screen was never inactive for longer than 48 hours.

In total, there are nearly five million screenshots in my sample, representing 6,815 hours of active smartphone screentime, and over four years of human experience in terms of subject-days. The average subject in this sample contributed 42,670 screenshots to the sample, representing 59.3 hours of screentime with a daily average of 4.6 hours. Most subjects in the sample contributed 14 days of screen activity. The extremes of daily smartphone use, with 0.1 and 17.7 hours per day respectively, represent extremely different approaches to screen usage and the role of smartphones in subjects' lives. Median active screentime of four hours per day is higher than the upper quartile daily estimates captured from a larger sample four years earlier (Christensen et al., 2016), potentially owing to the sampling pool of frequent smartphone users; possible sampling bias implications are discussed further in Section 9.

Table 3 summarizes eleven sample demographics within the available data. These features are not used as moderators in the main analysis but are provided here to inform readers as to the representativeness of the sample. Demographic information for the sample is limited, with 48 out of 115 subjects contributing valid self-report data available from the pre-survey. Missingness is assumed to be uncorrelated with demographic features. Further information on demographic data collection is provided in Appendix C. Subjects are well-distributed along multiple demographic features, as is expected by the quota sampling system used. Subjects are split nearly evenly along gender; no subjects reported nonbinary gender. Age is well-distributed between age 21 and age 68, with a median age of 44. Most subjects have at least some higher education past a high school diploma. Just over half of the sample is part of a married couple (or living as married), with most of the rest being single. Subjects are 69% white, similar to census representation. Roughly one quarter of subjects have household incomes less than $25,000 per year; 29% have household incomes between $25,000 and $50,000 per year. One quarter have household incomes between $50,000 and $100,000 per year, and 17% have household incomes higher than $100,000 per year. The vast majority of subjects speak English in the home. Politically, 17% of subjects state that they are not much interested in politics, while 44% state that they are very much interested in politics, with the remaining 40% being somewhat interested. Democrats and Republicans make up 44% and 42% of the sample, respectively. The cross-partisan feeling thermometer captures the feeling of self-reported Republicans toward Democrats, and vice versa; nonpartisans responded based on a preceding forced-choice partisan-lean question. Cross-partisan thermometer values range as low as 0, and 75% are 27 or less out of 100. Co-partisan feeling thermometers skew much higher, as would be expected. In sum, demographic information suggests that the sample provides basic representativeness per the American population.

Variable	Key	Value	Variable	Key	Value
Gender (%)	Female	56%	Race (%)	White	69%
	Male	44%		Black	14%
Age (years)	Minimum	21		Hispanic or Latino/a	10%
	25th percentile	34		Asian	2%
	Median	44		Multi/other	4%
	75th percentile	59	Household Income (%)	<$25k	27%
	Maximum	68		[$25k, $50k)	29%
Education (%)	HS or less	27%		[$50k, $100k)	25%
	Some college	35%		≥$100k+	17%
	Bachelor	25%	English at home (%)	Always	88%
	Graduate	13%		Mostly	8%
Marital status (%)	Single	30%		About half	4%
	Married	52%	Partisanship (%)	Democrat	44%
	Divorced	18%		Independent	8%
Cross-partisan thermometer (0 to 100)	Minimum	0		Republican	42%
	25th percentile	11		No Preference	6%
	Median	23	Attention to politics (%)	Not much interested	17%
	75th percentile	27		Somewhat interested	40%
	Maximum	39		Very much interested	44%
Cross-partisan thermometer (0 to 100)	Minimum	20			
	25th percentile	57			
	Median	74			
	75th percentile	85			
	Maximum	100			

Table 3. Demographic features for 48 out of 115 subjects. Information is self-reported. Percentages that do not add to 100 are affected by rounding. See Appendix C for further information.

In the next section, I describe the variables I have extracted from the raw sample data, and their connection to research questions in this dissertation.

Section 8. Variable Extraction

In this section, I describe the variables I have extracted from my raw data, and their connection to research questions in this dissertation.

Identifying screenshots containing political content

First, I identify political content in each screenshot. Hereafter, *political content* refers to any content containing any discussion of the following topic areas (Bode, 2016; Matthes et al., 2020): the U.S. presidential administration; politicians, including prospective, elected, or appointed government officials; political satire and political satirists; partisan groups, including political parties and think tanks; elections and election campaigns (at any level from national to local); policy debates and/or decisions (economic policy, foreign policy, social policy, other policy); political ideology and/or partisanship e.g., liberalism, conservativism, socialism, progressivism, fascism, populism, libertarianism, communism, republicanism or democratism; hate speech, racial inequity, identity threats or representation; terrorism, current American wars, US foreign policy, foreign affairs; analysis of or reactions to contemporary political events. These topics are based on national, state, or municipal governance (e.g., legislation, ideology), relevant system roles (e.g., Presidents, candidates, pundits), and matters of social justice (e.g., racism, terrorism) without accounting for source, format, or factuality.

Since classifying all 4.9 million screenshots in my sample by hand is impractical, I employ automated classification methods. To do so, I begin with a smaller subset of screenshots which was manually labelled and use that subset to train a supervised classification model. My initial set of political content ground-truth consists of 125,473 screenshots captured by 69 subjects analyzed in this dissertation. These screenshots were manually tagged by seven research assistants ($\bar{\kappa} = 0.59$) using the coding rules described above.[7] Screenshot tagging proceeded in two rounds, one in early 2020 (interrupted by the onset of CoVID-19) and one in 2021 conducted entirely remotely. Screenshots were randomly selected, stratified by subject and time, by

[7] Cohen's κ calculated among taggers' pairwise overlapped screenshots and averaged. Note that Cohen's κ value is biased downward in heavily unbalanced data (Xu & Lorber, 2014).

weekday, time of day, and 30-minute blocks within time of day, with an additional oversample subset of high-likelihood political screenshots based on the presence of basic political keywords. Within the resultant ground truth set, 1.99% were identified as containing political content.[8]

I use this ground-truth set to train and measure the performance of binary classifiers of political screenshots in the broader sample. The majority of features associated with screenshots are not suited for an application-agnostic classifier. For example, inputting image data into a political content classifier would bias a model toward visual formats by which political content is more often consumed. A recent demonstration of this can be found in Muise, Lu, Pan, & Reeves (2022), wherein unsupervised screenshot clustering based on image data returned groups of visually similar screenshots, often from single applications. Training a classifier using foreground application data or temporal features would have a similarly theoretically untenable impact on classification. Instead, text alone was chosen as a mostly format-agnostic asset for classification of political content. While politics can feasibly be presented and accessed in any format, text information is most clearly associated with semantic meaning, and is a highly common medium of political content on the smartphone. In the ground truth set, just 59 political screenshots contained no words (operationalized as containing fewer than five characters), comprising 2.4% of political screenshots and 0.3% of all wordless screenshots in the ground truth sample.

I preprocess text features to optimize classification performance and limit the influence of artifacts. *Spellcheck* was applied to the text from each screenshot using the *hunspell* package (Ooms, 2020). This removed obvious OCR format-disclosing artifacts (e.g., logos) and other

[8] I manually audited this tagged data following collection and found 451 instances of false negatives hand-tags, i.e., cases in which politics appeared on a screenshot but the screenshot was not labelled as such. This was due primarily to political survey-taking behaviors being found ambiguous by coders. These false negatives are reflected in the initial IRR, but in all following analyses and estimates, these are corrected.

artifacts (e.g., reading of the smartphone's battery symbol) and corrected words to nearby suggestions (e.g., disseratior → dissertation). To select a best classification method, I compared four approaches applied to the ground truth data: (*1*) bag-of-words based on 153 initial manually chosen stems related to political content in 2019 (*BoW*) (El-Din, 2016), (*2*) a manually improved bag-of-words approach based on expansion and restriction of the initial stems (168 stems), (*3*) supervised machine learning based on combinations of presence or absence of 168 stems in screenshots (Al-Amrani, Lazaar, & El Kadiri, 2018) and (*4*) supervised machine learning based on pre-trained text embedding applied to all text extracted from screenshots (Rudkowsky et al., 2018). Initial word stems underlying approaches (1) through (3) were based on a manual review of top political topics being discussed in late 2019, the period of data collection, as accessed through Google searches of top news articles. This initial set was intended as a starting point; the occurrence of one or more words from this initial set in a single screenshot would result in that screenshot being flagged as containing political content.[9] To move from the initial stem set of approach (1) to the improved set of approach (2), I manually examined the text found in screenshots that were *false negatives* and *false positives* according to approach (1). To reduce the *false negative* rate of approach (2) relative to approach (1), I expanded the stem set with 15 additional stems; to reduce the *false positive* rate of approach (2) relative to approach (1), I identified stems that frequently resulted in *false positives* and required that they appear with other political stems within a single screenshot in order for that screenshot to be flagged as political. The final set of 168 stems are shown alphabetized in Table 4.

abortion	caucus.	feminis	jeffrey.epstein	mueller	pundit	swallwell
affirmative.action	caucuses	filibust	john.oliver	nasty.woman	putin	swing.vote

[9] Additionally, in methods (1) and (2), screenshots captured while a gaming application was detected on-screen were classified as non-political regardless of stem presence in an effort to limit false positives from war-like terms such as *campaign*.

alex.jones	chuck.schumer	fiorina	kamala	o.rourke	quid.quo.pro	syria
americans	climate.change	fox.and.friends	kellyanne	obama	rachel.maddow	Taliban
amendment	clinton	G7	keystone	obamacar	racis	tax
andrew.yang	congress	gabbard	klobuchar	ocasio.cortez	rally	terroris
asylum	conservativ	gavin.newsom	law.maker	ouse.majority	rand.paul	tillerson
ballot	constitution	george.zimmerman	lawmaker	partisan	recession	transphobic
barack	covfefe	gerrymand	legaliz	pelosi	reform.	trayvon.martin
battleground.state	debate	ginsburg	legisla	pence	representative	treason
beto	deep.state	giuliani	lewandowski	polariz	republican	trevor.noah
biden	deepstate	govern	liberal	politic	RNC	trump
bigly	democrat	gun.control	limbaugh	pompeo	roger.stone	tucker.carlson
bigot	deplorabl	hassan.minhaj	lindsey.graham	populis	sanctions	united.nation
black.lives.matter	deport	hate.crime	lobbyis	president	sanders	warren
blm	DHS	hate.speech	locker.room.talk	primary.election	schiff	white.house
booker	DNC	house.minority	maga.	pro.choice	senator	white.nationali
border.wall	dogwhistl	human.right	manafort	pro.life	shooting	white.supremac
breitbart	donald.trump	immigration	mandate	prochoice	shutdown	witchhunt
brexit	economy	impeach	medicaid	prolife	socialis	.CIA.
buttigieg	elected.official	impeachment	melania	prosecut	sondland	.DOJ.
campaign	exit.poll	incel	merkel	protest.	stephen.colbert	.election
capitol	extradit	incumbent	migrant	protests	supreme.court	.facis
castro	FBI	jeff.sessions	minorities	proud.boys	susan.collins	.GOP.

Table 4. A list of 168 political word stems used in classification approaches (2) and (3). In approach (2), the presence of these word stems in screenshots identified screenshots as containing political content. In approach (3), the presence or absence of each stem from a screenshot's text was used as an individual feature with which to predict whether or not a screenshot contains political content, based on a random forest model trained on the ground truth set. This set was created based on political news events occurring in late 2019, and manually updated based on classification performance in a basic bag of words approach within the original ground truth set. In the table, the presence of a period (.) indicates a space between characters.

For approach (3), I treated the presence or absence of each of the 168 stems as 168 individual binary features upon which to predict political content in screenshots, using a random forest model trained on the ground truth set. The random forest model was tuned across various values of *m*, or word features, ranging up to 29. Six decision trees were applied per-forest, and intra-model forest selection was trained under three-fold cross-validation and three repeats, optimized on the *F* metric. The *F* metric is a more effective performance metric than *accuracy* in highly unbalanced classification tasks (Weiss, 2004). To determine performance metrics and optimal *m* of the model, I used five-fold cross validation of the entire model within the ground-truth set.

After determining the F-optimizing m value of 25, I then trained the model on the entire ground truth set. As the output of the random forest model is probability of political content within screenshots ($p \in [0,1]$), I determined the optimal threshold for binarized output based on the F_1 score, resulting in an optimal cutoff value of $p \geq 0.085$ for positive cases. Further parameterization information for each approach is explained in Appendix H.

For classification approach (4), I made use of pre-trained text embedding, specifically the GloVe 400k-vocabulary × 300-dimension dictionary trained on *Wikipedia* and *Gigaword* (Pennington, Socher, & Manning 2014).[10] For each screenshot in the ground-truth set, I identified all terms separated by spaces. Then for each individual term, I located the pretrained 300-dimensional vector associated with each term and averaged all vectors together, forming the mean semantic vector of the screenshot. Terms not found in the set were omitted from analysis. I then performed a principal component analysis on the ground truth set's resultant 300 semantic vectors. I trimmed the dataset to 19 out of 300 principal components which together explained 98% of variance in the data. I then applied a random forest training and classification procedure to this dataset, exactly as was done for approach (3), except that the features used were the 19 semantic principal components rather than the presence or absence of politics stems. Performance measures for all four approaches are shown in Table 5. Random forest based on word stems is the optimal classification method, as highlighted in green.

	First-pass bag of words	**Improved bag of words**	**R.F. classification w/ political words**	**R.F. classification w/ pretrained embedding**
Parameters	Single stem	Targeted stem	m = 25 stems,	m = 19 P.C.s,

[10] Note that training a text embedding model on the ground-truth set, rather than using a pre-trained model, is an approach that would suffer the same theoretical challenges as using image data to classify political screenshots. Some applications produce misleading format-based keywords. For example, the word "inbox" may occur frequently in political screenshots, but its presence is driven by visual format rather than the presence of political content.

	occurrence	restriction	threshold = 0.085	threshold = 0.4175
Accuracy	0.97	0.99	0.99	0.99
Precision	0.39	0.75	0.86	0.77
Recall	0.812	0.726	0.72	0.37
F	0.53	0.74	0.78	0.50

Table 5. Accuracy measures of four methods for identifying political content in screenshots: basic bag of words using intuitive stems, manually improved bag of words, random forest supervised machine learning (R.F.) using random selections of the improved stem set, and random forest supervised machine learning (R.F.) using pre-trained text embedding with principal component analysis, using the GloVe 400k vocab x 300 dimension dictionary trained on Wikipedia and Gigaword (Pennington et al., 2014). Both R.F classification methods utilized five-fold cross-validation for performance metric calculation, and 3-fold cross validation for tuning the number of randomly selected variables, m. Resultant probability values were binarized according to the F_1-minimizing threshold. F is the most useful accuracy measure for heavily unbalanced sets. More details in Appendix H.

Identifying segments of political content exposure

The screenshot is the smallest unit available in the study; I now explain how these screenshots are grouped together in users' experience. The relevant terminology used in this dissertation is diagrammed in Figure 2. One screenshot represents roughly five seconds of screentime. References to durations of any sort are based on this basic conversion of screenshot to time. *Session* refers to a continuous bout of smartphone usage, indicated in this data by a temporally contiguous set of screenshots each captured just seconds apart. The entirety of a subject's recorded screen behavior is referred to as a *screenome* (Reeves, Robinson, & Ram, 2020).

A *political content encounter* is an event in which political content appears on-screen when it was not on-screen previously. This is measured by the presence of a screenshot containing political content that is preceded by inactivity or a screenshot that does not contain political content. A *segment* (of political content exposure) is the span of time over which political content stays on screen following an encounter. For example, a political content encounter followed by seven screenshots, each five seconds apart, represents a segment of political content exposure lasting 40 seconds (8 screenshots × 5 seconds). The number of segments is definitionally equal to the number of encounters. Segments can traverse multiple applications, but usually do not. Segments are bounded by temporal separation of at least one screenshot not identified as containing political content, or any gap in screen activity (i.e., a break between sessions). This method is based on current literature which acknowledges multiple methods for determining gaps in smartphone behavior, but also acknowledges that the value of more complex approaches is as yet unclear (Peng, Zhou, & Zhu, 2020; Peng & Zhu, 2020; Van Berkel et al, 2016; Zhu, Chen, Peng, Liu, Dai, 2018).

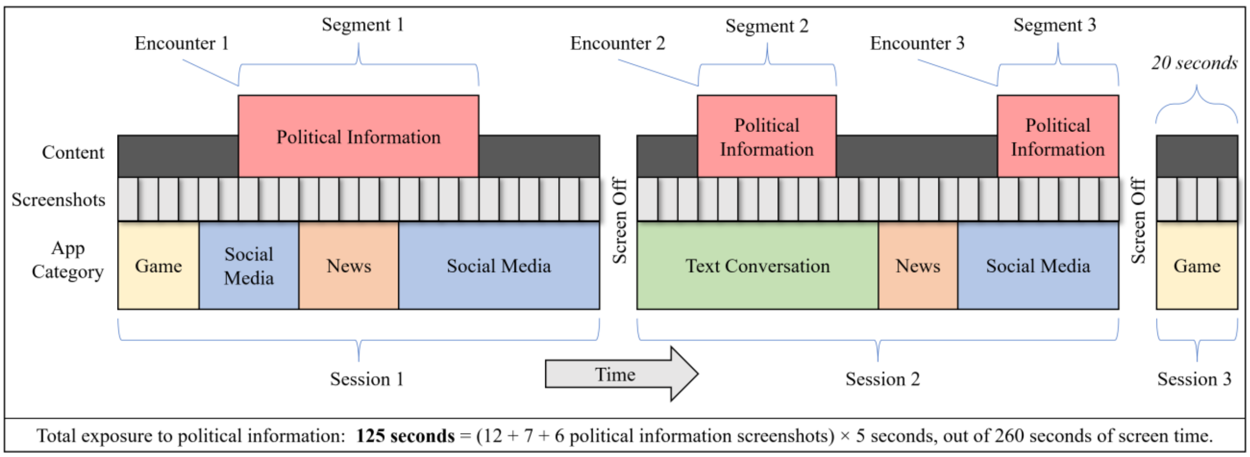


Figure 2. Diagram explaining terminology derived from or related to Screenomics data (Reeves et al., 2021). All terms are based conceptually upon smartphone usage, and their measures are all derived from underlying screenshot units taken at 5-second intervals. Screenshots are shown in gray. Sessions are sets of consecutive screenshots interrupted by gaps in screen activity of any length. The entirety of a subject's recording screen behavior is referred to as a screenome *(Reeves, Robinson, & Ram, 2020). Segments (of political content exposure) are initiated by encounters with political content and span any number of screenshots withing a session, independent of changes to the foreground application. Total exposure to political content is equal to 5 (seconds) times the number of screenshots containing political content over a given period.*

Note how all temporal metrics are derived from the five-second intervals at which screenshots are captured. Given this interval, there is variability in the duration over which political content was on-screen between screenshot captures. A single screenshot containing political content represents underlying exposure that may have lasted for $\overline{9.9}$ seconds (by having just barely been missed by the preceding and succeeding screenshot captures) or at minimum for

$0.0\overline{3}$ seconds, the standard refresh rate of smartphone screens. Assuming a uniform probability between these two extreme possibilities, the expected duration of such an exposure is five seconds in total ((5 seconds / 2) = 2.5 seconds before screenshot capture, (5 seconds / 2) = 2.5 seconds after screenshot capture). Durations of segments that span multiple screenshots are computed under the same assumption of uniformly distributed events between screenshots, i.e., 2.5 seconds preceding and succeeding the entire segment in expectation.

Throughout this analysis, I focus largely on the prevalence of segments lasting only five seconds, or the minimum detectable duration of political content exposure in my data. These short durations, while not instantaneous, are representative of extremely rapid exposure to politics in context of traditional assumptions of elaborative news consumption (Molyneux, 2018; Prior, 2009). As described in Section 5, this five-second duration is also a threshold beyond which recall and recognition of sequential on-screen advertisements declines to just 12%, suggesting a very low likelihood of elaboration (Goldstein, McAfee, & Suri, 2011). These rapid interactions with political content are interpreted as not allowing for meaningful cognitive engagement, in line with the presumption that longer durations are required for slower thinking (de Zúñiga et al., 2020; Kahneman, 2011).

Smartphone application categories

Regarding smartphone applications, in total, there were 2314 unique applications encountered by subjects in the sample. I group applications into 24 categories. My app categorization schema, available in full in Appendix D, is based upon application makers' self-declared categorizations on the Google Play Store, the default source for Android applications. I manually adjust some categorizations to better fit the goals of my dissertation. For example, given the preponderance of applications used by subjects to earn or gamble small amounts of cash, I have created a category 'survey/cash' which supplants the myriad labels of these applications created by individual app makers (e.g., "Lifestyle" for *Swagbucks*, "Social" for *YouGov*, and "Productivity" for *SurveyMonkey*). The primary use of this app classification system is to examine the degree to which political content is encountered via *news* applications or *social media* applications. The former category (e.g., the *CNN* app, the *Google News* app, or the *New York Times* app) is of interest given its historic role in conveying political content via other platforms, and its potentially waning relevance as described in Section 4. The latter category (including *Facebook*, *Instagram*, and *Reddit*) is of interest given the attention it has garnered from academics and the public alike, specifically as a common source of low-quality political content.

The category *Browser* (including *Chrome*, *Safari*, and *Edge*) can represent a variety of behaviors, given that any website with any content can be accessed via *URL* (including, e.g., *CNN.com* or *Facebook.com*). While I treat mobile browser applications as a standalone category, behavior on mobile browser applications essentially follows an unknown distribution across all other application categories. I proceed by assuming that behavior within the *Browser* category is not spent browsing a news website but may be political. This assumption is based on Allen et al. (2020), in which aggregated *ComScore* data is used to estimate that 2.78% of time that Americans spent on their smartphones from 2016 to 2018 was spent viewing news-formatted content via mobile browsers or applications, including fake news, partially political content, and non-political content.[11] Within that estimate, the proportion of news accessed via browsers versus applications (which includes a wide variety of aggregators and dedicated outlets) is

[11] See Allen et al. (2020) Supplementary Materials Table S6, rows four to six.

unknown, but given the ubiquity and comparative ease of using news applications, I assume that only a small minority of this 2.78% value is accessed via browsers.[12]

Section 9. Analysis & Results

In this section, I conduct my main analysis and discuss results, following the order in which I presented my research questions in Section 6.

Using the data described in Section 7 and variables described in Section 8, in this section I address the four research questions I posed in Section 6 of this dissertation, in order. That is, I will first describe the frequencies and durations of instances of political content exposure, with special attention to interindividual and intraindividual variation. I then examine how these frequencies and durations relate to application category (including news formats, and social media usage), and lastly how they relate to valence, arousal, and wordcount following recent research into mood management. To shed light on the idiosyncratic experience of each user, and to tie together diverse components of my analysis, I highlight five selected subjects in each group-level figure.

RQ1.1 What is the frequency with which political content appears on the smartphone screen?

To begin, I describe sample statistics regarding the frequency of political content throughout the entire sample. By frequency, I refer specifically to encounter frequency. Later in this section I describe durations. Political content appeared infrequently in the sample overall. Out of 4,907,091 screenshots captured across all 115 subjects, 92,988 contained political content.

[12] This assumption is bolstered by a finding in my analysis that 1.13% of screenshots are from *news* applications, on average across subjects.

This is equivalent to 1.89% of screenshots, or one minute and eight seconds out of every hour of screen activity. This value is similar to but lower than an estimate of 2.78% by Allen et al. (2020), which is based on news-format content rather than political content. The number of political content encounters (and equivalently, the number of segments) was 26,238 spread across 12,753 unique sessions. Next, I motivate my analysis by illustrating the complexity of political content exposure in the smartphone environment from a high level, using two figures. Figure 3 shows a histogram of encounter times throughout the day in the entire sample, broken into 15-minute blocks, standardized to each subject's local time. On the 24-hour clock (midnight to midnight), the average time at which an encounter with political content occurred was 13:16:08 local time, though encounters are broadly distributed through the standard waking hours and beyond. There is not a single block of 15 minutes in which a political content encounter did not occur. This picture contrasts sharply with older models of political communication founded on morning and evening primetime news consumption. In particular, the prevalence of political content encounters in the overnight and workday hours makes clear that political content is somewhat untethered to traditional divisions of the daily cycle.

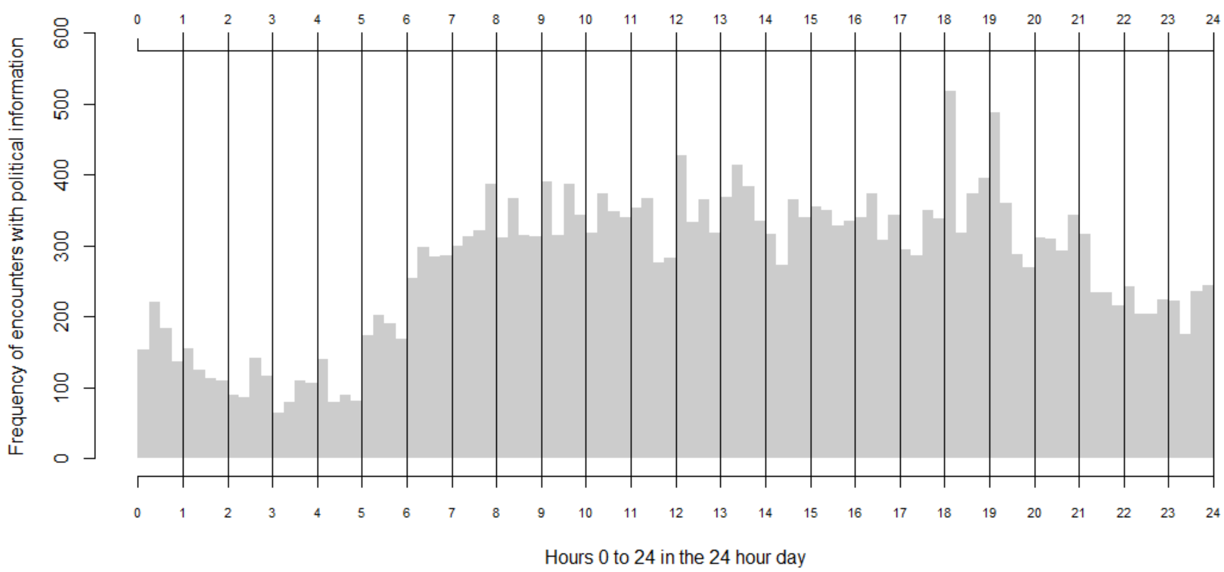


Figure 3. Political content encounters on the smartphone screen across 115 subjects (1463 subject-days) standardized to a single 24-hour clock, midnight to midnight. Encounters are bucketed into 15-minute blocks spanning the entire day. Note that 'encounter' refers to only the first screenshot in a series of continuous screenshots in which political content is identified, as explained in Figure 2.

Figure 3 shows variability throughout the day *in aggregate*. In that way, it masks variability both across and within individuals. In Figure 4, I provide an illustration of variability of political content exposure across subjects and across subject-days. Figure 4 depicts a week of smartphone screen activity, spanning 7 days (rows) for five subjects (columns). Each of the 35 individual panels shows a 24-hour clock (midnight to midnight) along a single horizontal timeline. Black coloration indicates smartphone screen activity of any kind; red coloration indicates the presence of political content on screen.

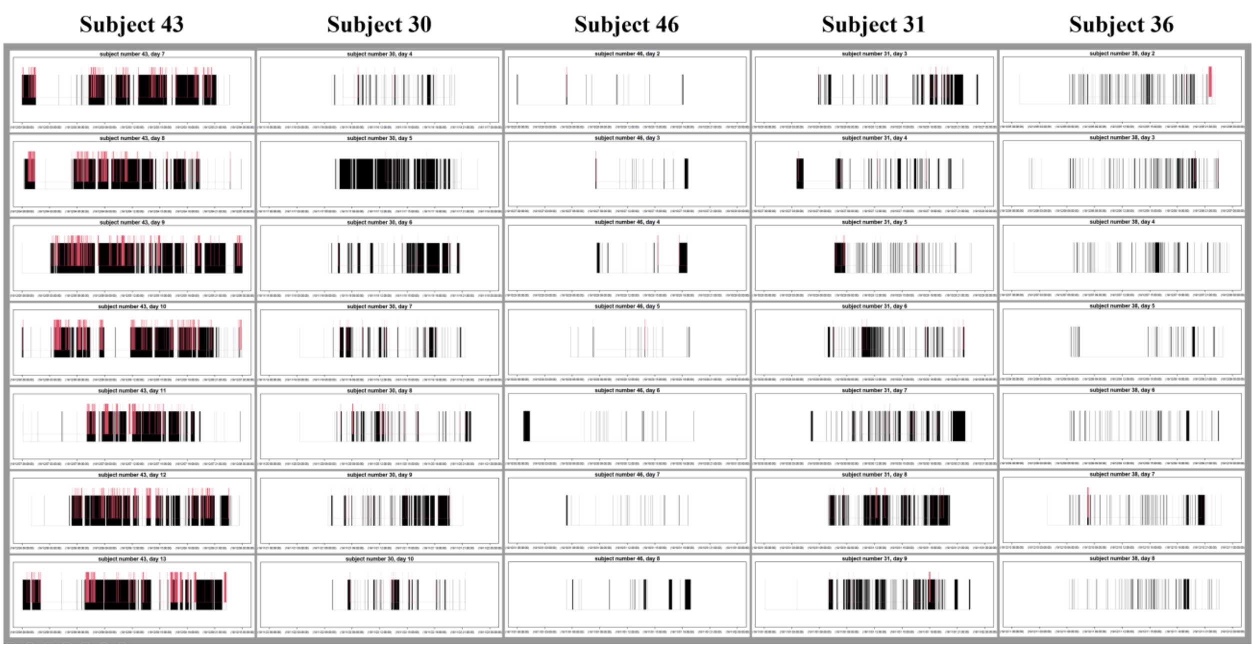


Figure 4. Visual illustration of inter-individual and intraindividual variability of political content exposure (red) in context of all screen activity (black) in 35 subject-days. Five columns represent five subjects; seven rows represent seven consecutive days for each. The leftmost subject, referred to as Subject 43, is an extremely frequent smartphone user and political content consumer, with active screentime spanning whole days (implying interrupted sleep) and political content consumed in bursts and fast instances across several applications at several hours. Other subjects in this figure follow a more typical distribution of social media usage and political content exposure, yet the specific sequencing of activity and inactivity in each subject-day is unique.

Across the five subjects (columns) of Figure 4, no two subjects share similar smartphone usage behavior, and within that behavior, they do not share similar political content exposure. The subject on the far left, referred to as Subject 43, is an extreme user of the smartphone and encounters more political content than any other subject (as is shown later in this section). Subject 43's behavior is organized into hours-long clusters of long-lasting sessions of usage. In contrast, Subjects 30 and 46 both use their smartphones much less frequently, with regular long breaks during the standard sleeping hours, and with very few encounters with political content. Subject 30 encounters political content several times in each day, yet the resultant segments are so short that they are difficult to discern without magnification. The remaining subjects are each unique with clearly distinct features to both their overall usage and political content encounter frequency.

Perhaps more interesting is the degree of variation *within* individuals. For example, at the top right corner, we see that Subject 36 experienced an uncommonly long segment of political content exposure before sleeping, lasting longer than all segments in the next four days combined. Subjects 30 and 31 both show heavily fluctuating screentime across consecutive days. With increased granularity in time — days to hours to minutes — the likelihood of political content being consumed with any particular frequency is increasingly unpredictable in any particular unit of time, a point I will address later in this section. Appendix B contains a link to every subject-day in my sample, following the style used in Figure 4. Motivated by the illustration in Figure 4, the remainder of my analysis addresses variation across and within subjects.

RQ1.2 How does political content exposure frequency vary across individuals?

In this subsection, I demonstrate inter-individual variation in exposure to political content. In Figure 5 I rank each subject according to political content exposure, measured in terms of daily seconds (Panel A), unique encounters (Panel B), and percentage of daily screentime (Panel C). Subjects are organized according to their ranking in Panel A, and the ranking is held constant in Panels B & C such that the same subject occurs in the same horizontal level. The five subjects shown for illustration in Figure 4 are highlighted here as well for additional clarity.

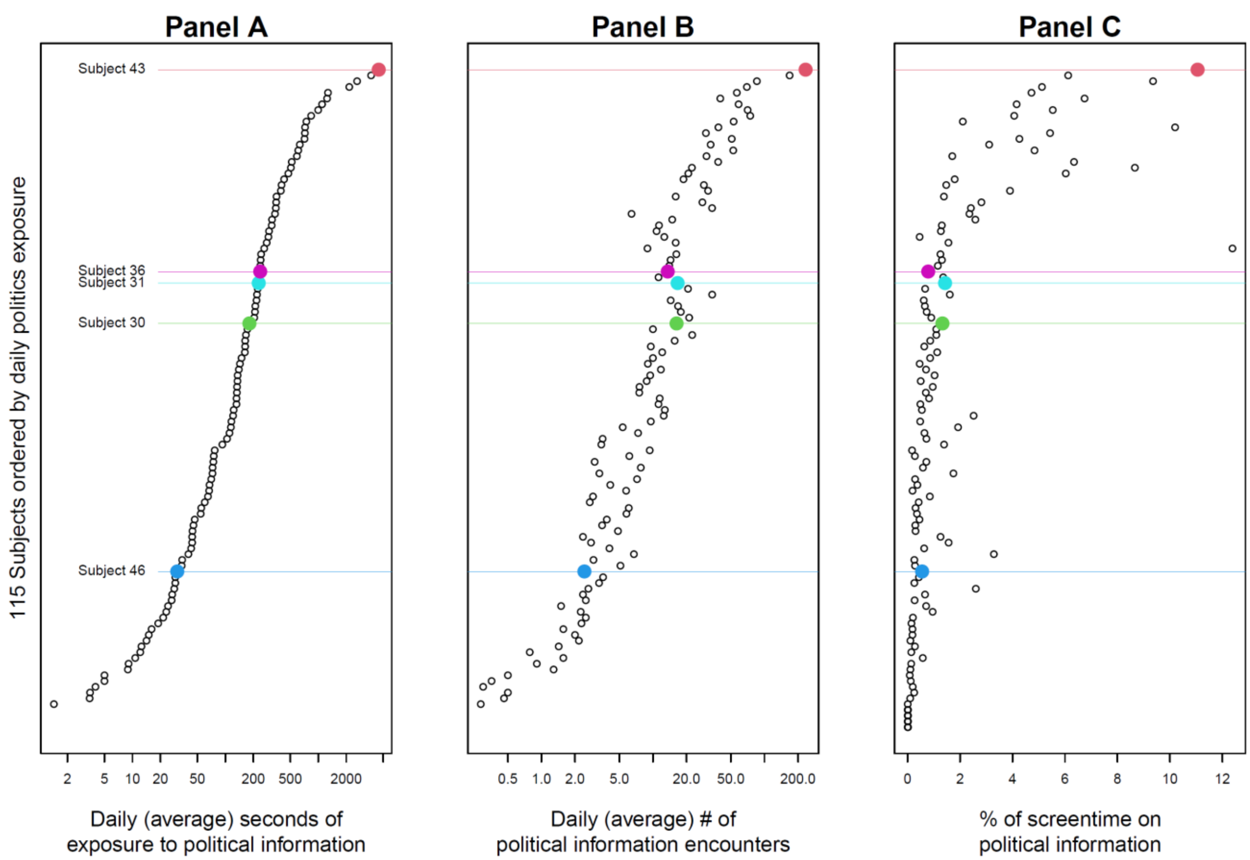


Figure 5. Inter-individual variation in exposure to political content in terms of daily seconds (Panel A), percentage of overall daily screentime (Panel B), and daily frequency of unique encounters (Panel C). Daily values are averages calculated across all unmanipulated subject-days, per subject. In Panel A, subjects are ordered according

to their x-axis value; in Panels B and C, subjects maintain their ordering from Panel A. Note the log scales on Panels A and C, and the corresponding subjects shown from Figure 4.

The most prominent finding is that most political content exposure is driven by relatively few key subjects, with most subjects consuming negligible amounts of political content through their smartphone. Recall that political content exposure comprised 1.89% of all screentime in the sample. In Panel C, we see that Subject 31 lies very close to this group-level average, and many subjects are lower, with four subjects encountering no political content. However, those with higher percentages range far beyond and exhibit a markedly different experience of political content exposure. Grouping these two broad groups of subjects together within single group-level statistics does worse than lose information about the sample. Research that relies on averaging political content exposure across this starkly varied distribution assigns political content into diets that contain virtually none, and thus can overestimate the level of political interest in the population.

To statistically gauge the utility of group-level summarization in this sample, I use the *coefficient of variation* or *CoV* (Bedeian & Mossholder, 2000). The *CoV* of a random variable is equal to its standard deviation divided by its mean, thus providing a standardized measure of how *incorrect* a group mean is for estimating the value of a single point within the group, i.e., a standardized measure of the magnitude of the ecological fallacy if enacted on a given sample. Across the 115 subjects, the average daily number of political content encounters (Panel B) is 18, and the *CoV* is 1.7. This means that if the sample mean of 18 were used to estimate the subject-level value of any single subject, the estimate would be off by a factor of 1.7 in expectation. Turning to Panel C: the *CoV* of political screentime percentage is 1.4; this suggests that, in expectation, a sample-level estimate of time dedicated to political content is off by a factor of 1.4

from any one subject's lived experience. I make further use of these *CoV* values throughout my analysis.

Cross-sectional aggregation also sidesteps the importance of individuals driving virtually all political content consumption. Consider for example the uniqueness of the top political content consumer, Subject 43, shown in red. Subject 43 consumed 75 minutes of political content on an average day, comprising 11% of their overall daily screentime on the smartphone. In contrast, the 70 subjects least exposed to political content collectively saw 74 minutes of political content on an average day, despite collectively experiencing 319 times the overall screentime as Subject 43. In terms of unique encounters with political content, Subject 43 encountered political content 234 times on an average day, while the median subject experience was just 9 per day. While Subject 43 is clearly an extreme case relative to the rest of the sample, their individual experience is equally genuine as that of the modal subjects, and perhaps of much greater interest and relevance to political communication research. Regarding the centrality of timescales in political communication research, most subjects' daily exposure to political communication is so slight that their modal experience of political content exposure is not likely to last long enough for longer elaboration processes.

To summarize this subsection: (1) most exposure to political content is driven by a small share of subjects, suggesting that group-level means describe inter-modal and potentially non-existent experiences; (2) the highest percentage of screentime in which political content was on screen was 11%, while the median percentage was 2%; (3) group level estimates of political content exposure in this sample are incorrect by a factor of 1.7 in expectation when applied to a single subject. In the next subsection, I show how the variation across individuals compares to variation *within* individuals over time.

RQ1.3 How does political content exposure frequency vary within individuals?

Drawing on the times-persons-variables cubes of Figure 1, I now rotate from discussing variation across *persons* to discussing variation through units of *time* within individuals. To directly compare these two types of variation, I make further use of the *coefficient of variation* and loosely adapt an approach used by Fisher et al. (2018).

In the preceding subsection, each subject's exposure to political content (measured in three ways in Figure 5) was presented as a single point statistic. As was illustrated in Figure 4, however, subjects' behavior is not static or uniform over time and thus not well captured by any one value estimate. This is increasingly apparent at smaller and smaller timescales. To summarize this point statistically, I calculate the within-subject *CoV* of two core measures of political content exposure: the frequency of political content encounters and the percent of screentime spent on political content. The results of this analysis are shown in Figure 6. In either panel the *y*-axis shows the *CoV* of the measure variables, and the *x*-axis is divided into 4 temporal levels of aggregation. At the origin, I estimate the *CoV* of variable estimates using each subject's entire data collection period. The *CoV* is marked as 0, as this is a single point estimate for each subject. At the remaining three temporal levels, I estimate each subject's intraindividual *CoV* at the corresponding temporal level by bucketing the subject's data timeline into a vector of discrete temporal buckets. Five subjects are highlighted in colors to provide continuity from previous figures; all 115 subjects are shown in the background in gray, and sample averages are calculated at each point in bold black.

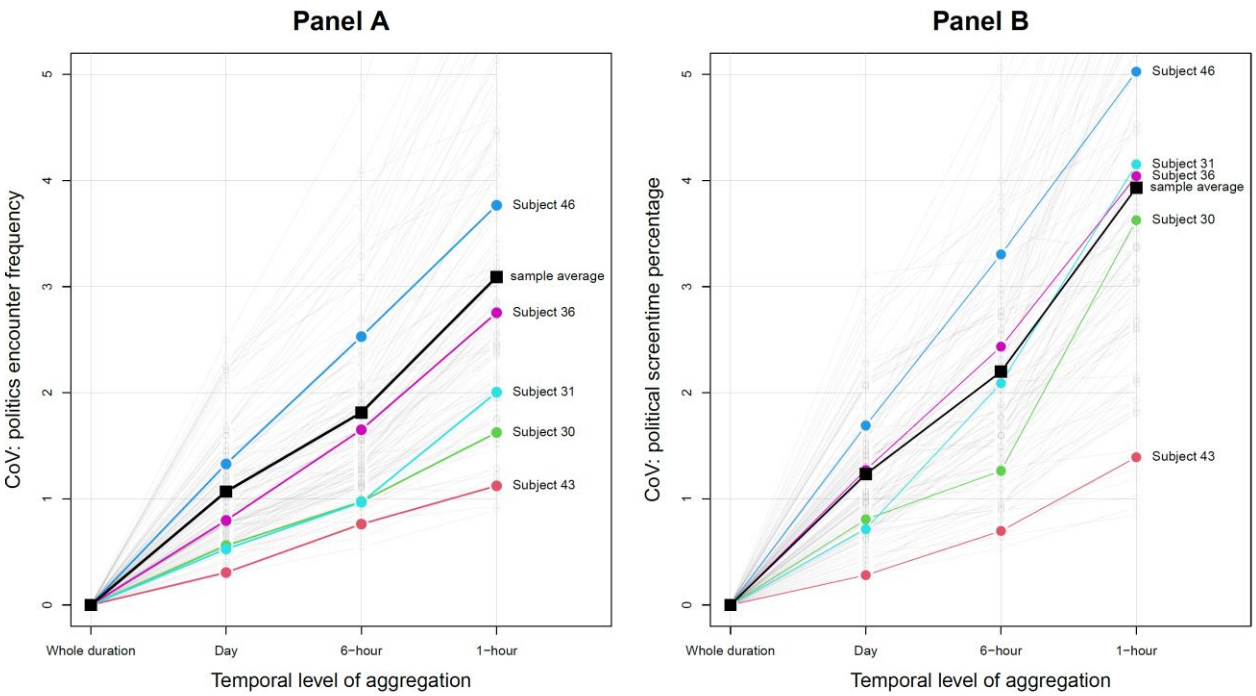


Figure 6. *Coefficient of Variation (CoV) of daily frequency of political content encounters (Panel A) and the percentage of screentime during which political content was on-screen (Panel B). The y-axis shows CoV values ranging from 0 to 5, and the x-axis shows temporal levels of aggregation ranging from a subject's entire study duration to hour-long blocks at the most granular. All 115 subjects are shown in gray, with five subjects highlighted for reference, and sample-wide averages shown in bold black squares. On average across the sample, estimates of hourly politics encounter frequency misrepresent a randomly selected hour by a factor of 3.1, in expectation. On average across the sample, estimates of political screentime percentage misrepresent a randomly selected hour by a factor of 4, in expectation.*

By calculating across time, the *y*-axis values indicate how inaccurate a subject's average behavior would be in predicting the behavior of the subject at a given point in time. For example, if a researcher calculated Subject 46's hourly number of encounters with political content (Panel A), that estimate would be off by a factor of 3.72 relative any single hour in the sample, in expectation. Across all subjects in this sample, the average *CoV* of political content encounter frequency is 1 at the daily level, 1.8 at the level of 6-hours, and greater than three at the level of hours. For political screentime percentage, these numbers are even higher, and similarly increasingly with increased temporal granularity. For some subjects, the *CoV* for either value ranges as high as 10, meaning that an hourly estimate would be off by an order of magnitude in expectation.

Figure 6 shows that intraindividual variation is on par with interindividual variation. A cross-sectional average of daily political content encounters is incorrect by a factor of 1.7 when applied to individuals in expectation (from preceding subsection). At the level of six-hour blocks, interindividual variation and intraindividual variation are nearly equal, and at finer granularity, intraindividual variation is greater than inter-individual variation. More broadly, Figure 6 helps to validate the role of timescale explication in measurement. A single core variable, measured in a single person, displays monotonically increasing variability at successive levels of aggregation.[13]

To summarize this subsection: (1) subjects' intra-individual variation in political content exposure — measured as unique encounters or as percent of screentime — is on par with inter-individual variation; (2) on average, a subject's daily political content exposure is expected to be inaccurate by a factor of 4 in estimating any particular hour of the subject's experience; (3) intraindividual estimates are increasingly inaccurate when applied at short time spans, tracing the value of time scale explication in measurement. In the next subsection, I transition from analyzing exposure *frequency* toward *durations* of political content segments.

RQ2.1 What is the distribution of durations of political content segments on the smartphone?

[13] The point is echoed by Muise et al. (2022a), wherein the estimation of an intraindividual news diet, and thus assignment of partisan bias to individuals, is shown to depend on the temporal level of aggregation. See Supplementary Materials, Section 9.

As described in Section 5, the smartphone enables incredibly fast interactions with any topic at any time, including political content. Whereas many studies focus on the frequency of encounters with political content, the length of time that political content is engaged with following a single encounter is unbounded on the smartphone: an instant or a calendar day are both feasible, unlike thirty-minute blocks of content on television. Here, I analyze how long political content segments typically last on the smartphone. Drawing on Goldstein, McAfee, & Suri (2011), I focus in particular on segments lasting just five seconds: the minimum detectable duration in my sample and a threshold at which recognition & recall rate of advertisements is as low as 12%.

To begin, I tabulate all political content segments in the sample and plot their distribution, as shown in Figure 7. Political content segment durations follow a power law distribution, much like durations of *sessions* of smartphone usage (Muise et al., 2022b) and segments of app usage (Rula, Jun, & Bustamante, 2015). In bold black is the distribution of segment durations for the entire sample of segments, pooled together. In low opacity around it are the distributions of segment durations for each of the 111 subjects who encountered political content. As with prior figures, I have highlighted five subjects for continuity and clarity. In the entire sample, 44% percent of political content segments lasted just five seconds, and 61% lasted ten seconds or less. Only 16.97% lasted longer than thirty seconds. As a robustness check on this result, I point out that the 26,238 unique segments detected in the sample occurred across 12,753 unique sessions.[14] The large number of unique sessions in which segments were detected implies

[14] To clarify, a *segment* is a bout of continuous exposure to political content, proxied by consecutive screenshots flagged as containing political keywords by a keyword-based random forest classifier. A *session* is a bout of continuous screen activity, proxied by consecutive screenshots separated by inactivity in time. See Figure 2.

that the shortness of segments is not reflective of longer latent segments split up into chains of short segments by my detection strategy.

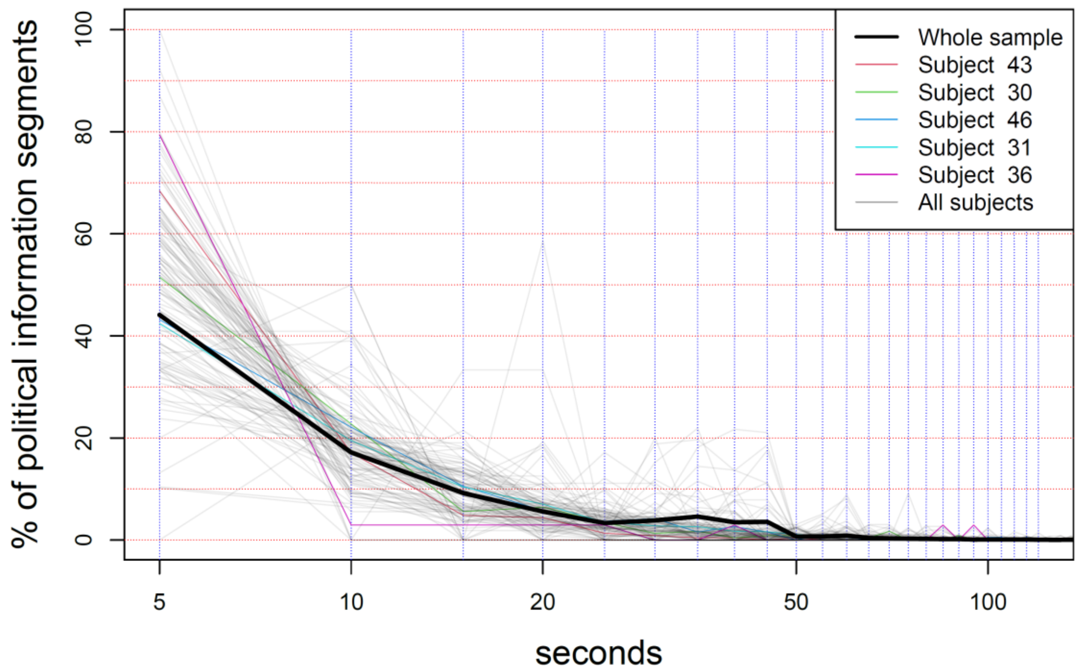


Figure 7. Distribution of political content segments across duration lengths. Segment length distribution for all segments in the data is shown in bold; each subject's individual distribution is underlaid in thin lines to shown inter-individual variation. Forty-four percent of all political content segments lasted just five seconds, the minimum detectable duration in the sample, and a duration at which recall and recognition rate is as low as 12% just minutes after exposure (Goldstein, McAfee, & Suri, 2011).

Figure 7 makes clear that the pace of the smartphone experience is in clear contrast to the news formats that formed the basis of conventional measurement strategies in political communication. An extremely small proportion (< 1%) of political content encounters lead to segments lasting one minute or longer; the preponderance of encounters with political content on the smartphone lead to extremely short-lived segments. While the impact of these short segments is not directly measured in my data, inductive logic allows for estimation of how they impact estimates of overall exposure. Out of the 26,238 political content segments detected in the sample, 11,545 lasted just five seconds (44%). These one-off encounters represent 12% of the 92,988 political screenshots detected, or equivalently, 12% of all time exposed to political content. An additional 17% of 26,238 encounters lasted ten seconds, representing 9.5% of the 92,988 political screenshots in the sample. Estimated recall and recognition rates in sequence are 12% and 14% for screen-based advertising lasting five and ten seconds respectively, measured a few minutes after exposure; see Table 1. Thus, the majority of political content encounters occurred in a manner that does not lead to information retention, comprising more than a fifth of all political content exposure in terms of time on screen. Considering the very low likelihood that these instances of exposure lead to elaboration, processing, or opinion formation (based on information retention rates alone), their contribution to political learning or socialization is unclear apart from the invocation of fast-thinking processes, i.e., in the realm of the emotional and psychological, if having any impact at all. Moreover, given the broad temporal spread of segments across sessions, any quick impressions left by fast segments are apparently unlikely to compound by occurring together in bursts.

To summarize this subsection: (1) nearly half of segments last only the shortest possible detectable duration, five seconds; (2) the expected recall and recognition rate of the content of *most* segments is under 15%, based on advertising research (3) only a negligible amount of political content segments last longer than one minute. In the next subsection, I examine interindividual variation in subjects' distributions of political content segment durations.

RQ2.2 How do political content segment durations vary across individuals?

As shown in the background of Figure 7, each subjects' experience of political content on the smartphone is similarly fast-paced. However, there are crucial differences across subjects that are telling and meaningful for the study of political communication. In Figure 8 I plot the mean, median, and maximum segment duration for all 115 subjects in Panels A, B, and C respectively.

Subjects in Figure 8 are ordered according to mean segment duration in Panel A. Five subjects are highlighted for clarity and continuity with prior figures. Not shown, the minimum segment duration for all but one subject is 5 seconds or none at all.

All subjects' mean and median segment duration was under one minute long, and the majority of subjects' median segment duration was five seconds. For the majority subjects, the longest political segment was shorter than three minutes, and for many subjects the longest segment was under one minute (Panel C). The longest single segment in the sample lasted 29.7 minutes. This was a midnight viewing of the *Joe Rogan Experience* via *YouTube*. While a popular source for commentary, Rogan is not held to the same journalistic standards as the traditional news outlets and has been criticized for popularizing false information (Rozner, 2018). That he drives the longest-lasting political segment in the sample is telling of how the smartphone environment is extremely different from the traditional media channels upon which our field is based. Subject 43, the subject which was found to experience the most encounters with political content and the most political content exposure overall, ranks 26th out of 115 subjects in terms of mean segment duration and has a median segment duration of ten seconds. This finding contravenes an intuitive assumption that short bouts of engagement with political content are a reflection of *disinterest* in political content.

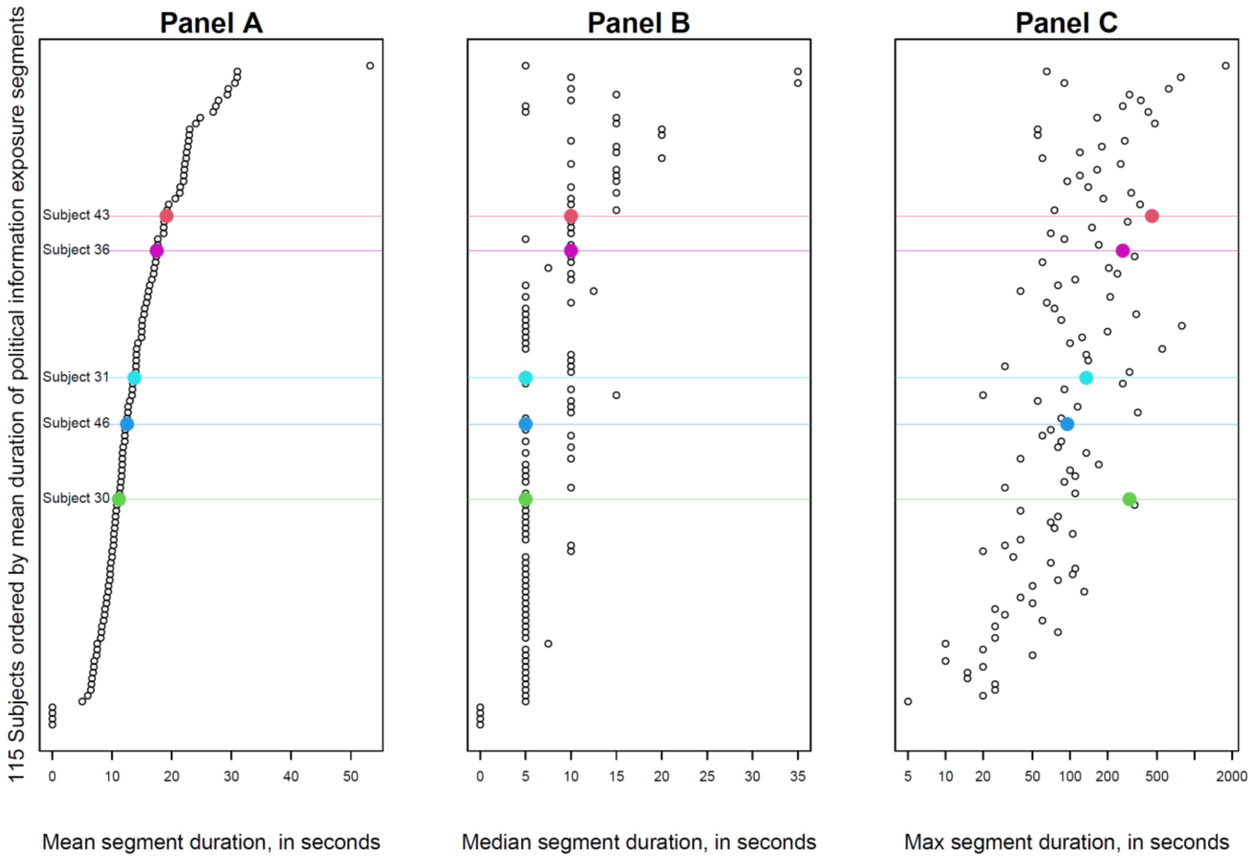


Figure 8. Mean duration (Panel A), median duration (Panel B), and maximum duration (Panel C) of political content segments for 115 subjects. Maximum duration is presented on a logarithmic scale. Subjects are ordered according to mean segment duration. Five subjects are highlighted for clarity. Most subjects' median segment duration was five seconds.

The apparent incongruity between Subject 43's short median segment duration and high degree of overall political content exposure motivates a look each subject's intraindividual experience of political content segments relative to their political content encounters. I address this in Figure 9. In Panel A, I show the percent of each subject's political content encounters that lead to segments lasting five-, ten-, and fifteen-seconds respectively. White space to the right of each subject's bar (up to 100 on the *x*-axis) represents the percent of that subject's political content encounters that led to segments lasting longer than fifteen seconds. For all but eight out of 115 subjects, the strict majority of political content segments lasted just 15 seconds or less; an additional four out of 115 subjects never encountered political content. As is expected in a fast-paced and low deliberation setting, there is no relationship between a subject's percentage of five-, ten-, and fifteen-second segments.

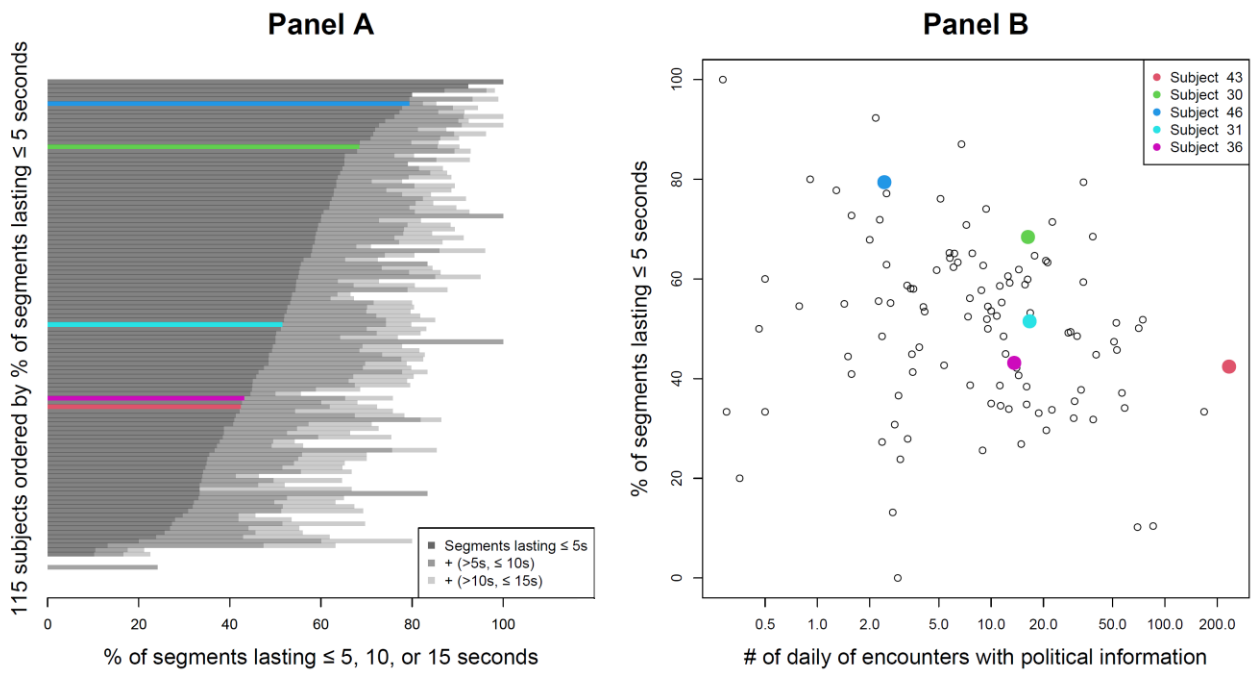


Figure 9. Subject-level segment durations. Panel A arranges 115 subjects according to the percent of their political content encounters that lead to 5-second segments. In progressively lighter shading, I show the additional percentage of encounters leading to 10-second-long and 15-second-long segments. In Panel B, I plot each subject according to the percent of their political content segments lasting five seconds (y-axis) and their number of daily encounters with political content. Note the logarithmic scale used in Panel B. Five subjects are highlighted for clarity.

In Panel B, I plot each subject according to their number of daily encounters with political content (*x*-axis) and the percent of their political content segments lasting five seconds (*y*-axis). There is only a weak correlation (ρ = -0.14) between these two variables. Subjects who encounter politics less often, as exemplified by Subject 46 in blue, are somewhat more likely to exhibit a higher intraindividual percentage of five-second segments. Still, Subject 43 experiences five-second segments at just under the rate of the sample as a whole. To the extent that encounter-based measurement of political content exposure is upwardly biased with respect to underlying information experience (as was addressed in analysis of RQ1.1), Figure 9 shows that this bias is not uniform across subjects. For example, Subject 30 encounters political content about 30% more often than does Subject 36, yet 70% of Subject 30's encounters lead to segments that are five seconds long and thus likely not memorable, versus 44% for Subject 36.

To summarize this subsection: (1) the majority of subjects' median segment duration is just five seconds; (2) in the entire sample, the median segment duration is 35 seconds or less; (3) there is a very weak negative correlation between number of encounters with political content and the intraindividual percent of segments lasting just five seconds. In the next subsection, I examine subjects' intraindividual distributions of political content segment durations through time.

RQ2.3 How do political content segment durations vary within individuals?

Results presented throughout the preceding analysis have provided some depiction of intraindividual experience of political content segments. Figure 3 showed that political content encounters occur across the entire 24-hour day, concentrated in the waking hours. Figure 4 showed that subjects' day-to-day and hour-to-hour political content exposure is highly varied, and Figure 6 quantified that this variation is on par with or in excess of interindividual variation. Lastly, Figure 9 showed that nearly all subjects' encounters primarily lead to segments lasting just 5 seconds. In this section, I build on these findings to quantify the structure of intraindividual variation in political content segment durations. First, I apply the *CoV* measurement strategy to compare intraindividual and interindividual political segment duration measurements. Second, I examine the relationship of segment durations to encounter frequency and exposure time, and the ergodicity of these bivariate relationships. Third, I examine whether or not intraindividual political content segment durations are related to overall session durations, i.e., if temporal screen activity patterns correspond to temporal segment patterns.

As explained earlier in this section, the *CoV* is a standardized measure of how accurately a group average describes units within the group it summarizes, where units can refer to persons or occasions. To compare intraindividual and interindividual variation in segment durations, I first calculate the *CoV* of three interindividual measures of subject-level segment duration tendency: mean segment duration, median segment duration, and maximum segment duration, as was shown in Panels A, B, and C of Figure 8 respectively. The *CoV* of mean segment duration is 0.52, the *CoV* of median segment duration is 0.66, and the *CoV* of maximum segment duration is 1.37. This means that the average subject-level mean segment duration is off by a factor of 0.52 in describing any individual subject's mean segment duration, with analogous interpretations for the median and maximum *CoV*. Next, I calculate the *CoV* for each subject's intraindividual distribution of segment durations and take the average of this value across the sample. The resulting average intraindividual *CoV* is 1.2, more than double the interindividual *CoV* for the most comparable metric, mean segment duration. In essence, a typical subject's average experience of political content segment duration is more consistent with that of other individuals than with their own intraindividual experience. This finding demonstrates that singular summary values of individuals may provide a deceptively flawed view of political audiences.[15]

Thus far in my analysis I have strictly compared univariate interindividual variance with univariate intraindividual variance through application of the standardized *CoV* metric. In doing

[15] Muise et al. (2022a) demonstrated how this kind of intraindividual variation can undermine seemingly significant findings calculated from flattened inter-personal estimates. At the group level and cross-sectionally, a large swathe of Americans appear to consume highly skewed partisan news diets. However, by examining intraindividual variation, they demonstrated that the vast majority of apparent partisans were shifting their news diet month-on-month, undermining speculation about large echo chambers. See Figure 2.

so I have demonstrated that researchers can easily commit an ecological fallacy by not only applying group means to individuals, but also by applying subject means to subjects at all time points, and risk similar levels of inaccuracy in either action. Next, I compare interindividual bivariate relationships with intraindividual bivariate relationships. A bivariate relationship that is similar both interindividually and intraindividually (i.e., with the same polarity) is *ergodic*, while a bivariate relationship that is dissimilar interindividually versus intraindividually is non-*ergodic*.

In Figure 10 I plot the pairwise Pearson correlations of segment duration, encounter frequency, and exposure time both interindividually and intraindividually. To do so, I first calculate each variable's value for each subject-day, and then calculate the intraindividual correlation of these day-level variables for each subject, producing a single correlation value for each variable pair. Separately, I calculate each subject's three average variable values through their entire data collection period and then calculate the interindividual correlation between those averages. The result of this process is two bivariate correlation statistics for each variable pair: the average intraindividual correlation and the interindividual correlation of average behavior. The results are displayed in Figure 10, along with identical ergodicity analyses conducted at the *hourly* level rather than the daily level. Days or hours in which subjects encountered no political content are treated as missing values. All variables are standardized to mean 0 and standard deviation 1 immediately prior to correlation calculation, interindividually or intraindividually.

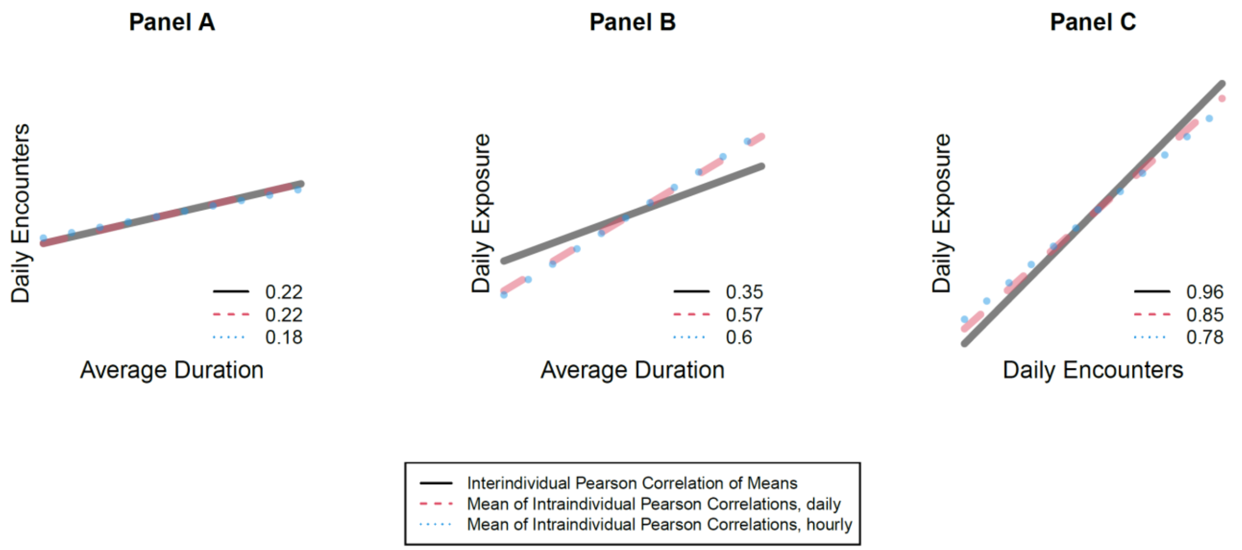


Figure 10. Interindividual versus intraindividual pairwise correlations of three variables summarizing political content exposure: average segment duration, number of encounters, and total exposure. Interindividual correlations are calculated from the subjects' average behavior in the data collection period. Intraindividual values are the average of intra-subject correlations between variables across days. Correlations are Pearson method, with standardization immediately prior to estimation. Correlations with opposing polarity represent non-ergodic pairwise relationships; all relationships display ergodicity.

In all three cases, and at either temporal granularity, the relationship between each variable pair is ergodic. This is clear from the identical polarity and similar magnitudes of correlations calculated interindividually and intraindividually. In Figure 10 Panel A, I show that the correlation between average segment duration and daily encounter frequency is 0.22 both interindividually and intraindividually at the daily level, and 0.18 at the hourly level. This means individuals with longer average segment durations tend to have more daily encounters, but also that for a given subject, on days (and hours) with more encounters than usual, the subject tends to experience longer segment durations than usual. This relationship is weak in magnitude in that the correlation is only 0.22, but it is striking in its robustness to level of analysis.

In Panel B, I compare average segment duration with daily exposure time. Interindividually, the correlation between these two variables (as subject averages) is 0.35. Intraindividually, however, the correlation is 0.57 at the daily level, and 0.6 at the hourly level.

While this relationship is ergodic, it is clearer intraindividually than interindividually. This is intuitive: on days in which subjects experience a higher average segment duration than usual, they also experience a higher-than-average amount of exposure time. At the hourly level, single segments comprise a larger share of overall exposure time within the hour; at lower granularities, this intraindividual correlation value is expected to increase further. In Panel C, I show the relationship between daily encounter frequency and daily exposure. Interindividually, these variables are almost perfectly correlated: subjects with the highest daily encounter frequency have the most overall exposure. For a given subject, in a day (or hour) in which a subject encounters political content more frequently than is typical for them, exposure time is also higher than is typical. This intraindividual relationship weakens with decreases in granularity, as fewer encounters can fit into increasingly small timespans. Altogether, the interindividual relationships found in each pairwise comparison confirm intuition, and the intraindividual relationships are extremely similar to their interindividual counterparts, suggesting that the core components of political content exposure in the smartphone environment are ergodic in time. That is, to the extent that cross-sectional relationships are used to describe or predict intraindividual change, such applications are appropriate within these core features of exposure.

Lastly, I examine how subjects' intraindividual distribution of political content segment durations compares with their intraindividual distribution of session durations.[16] A positive correspondence between these two variables would suggest that subjects' temporal engagement with political content is related to their intraindividual usage style. To conduct this analysis, I apply an approach from Muise et al. (2022b), using what the authors refer to as a *rapidity score*,

[16] Once again, a *segment* is a bout of continuous exposure to political content, proxied by consecutive screenshots flagged as containing political keywords by a keyword-based random forest classifier. A *session* is a bout of continuous screen activity, proxied by consecutive screenshots separated by inactivity in time. See the diagram in Figure 2 for clarification.

which is in-turn based on the power-law distribution of durations common to all subjects' session durations and segment durations. A rapidity score is the exponential degree which minimizes mean squared error (MSE) when fitting an exponential distribution to the distribution of session [segment] durations. This value summarizes the relationship between session [segment] duration and frequency. The higher the rapidity score, the greater the propensity for shorter durations relative to longer ones. As rapidity scores are dependent on duration distribution range, here I fix the range for both segments and sessions to [0, 20 minutes] discretized into five-second buckets.

In Figure 11, I compare the segment rapidity scores to session rapidity scores for all 111 subjects who encountered any political content. Along the *x*-axis, segment rapidity scores range from 0.05 to 10 along a logarithmic scale; along the *y*-axis, session rapidity scores range from 0.05 to 2.1, also along a logarithmic scale. Five familiar subjects are highlighted in color for clarity and continuity. Note that rapidity scores at low integer values imply well-known distributions. A rapidity score of zero, not found in any subject, represents a white noise pattern in which all session durations [segment durations] occur equally frequently. In particular, a rapidity score of one means that sessions [segments] follow a *pink noise* pattern, common in film scene durations (Cutting, DeLong, Nothelfer, 2010) and underlying how humans allocate attention (van Orden, Holden, & Turvey, 2003).

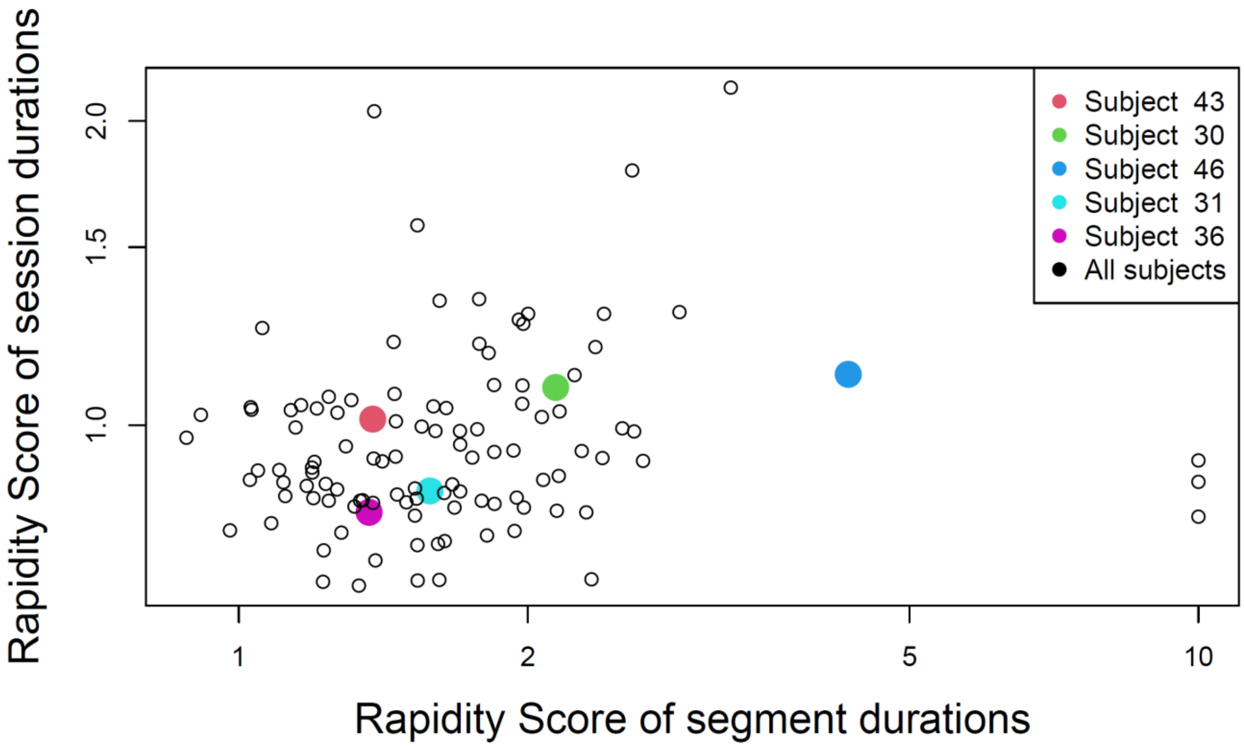


Figure 11. *Subject-level scatter plot of rapidity scores of segment durations against rapidity scores of political content segment durations. A rapidity score is an MSE-minimizing degree describing the relationship between event and duration, where duration = (freq)$^{-degree}$ and possible range is in [0, 20 minutes]. Higher rapidity scores implies greater tendency toward faster events versus longer ones. If a correlation exists between session rapidity and segment rapidity, then the lengths of subjects' segments might reflect underlying tendency toward longer durations of screen activity, or longer attention spans. However, there is no correlation. Subjects with no political content encounters are not included in this plot.*

In this sample, sessions of smartphone usage are largely distributed in a manner that resembles pink noise (rapidity score ≈ 1). That is, for most subjects, the likelihood of a smartphone session lasting for duration *f* is approximately inversely proportional to *f*. In contrast, segment rapidity scores tend to be higher. This is expected given that 44% of all segments last for the minimum duration in the data. The three extreme outliers along the *x*-axis, with segment rapidity score = 10, represent subjects with very few encounters with political content, and thus the overwhelming majority of their segment durations are just five seconds long. This phenomenon is also reflected in the segment rapidity score of Subject 46, shown in blue, who encountered the least political content out of the five highlighted subjects. Ultimately, the two sets of subject-level rapidity scores are not meaningfully correlated (ρ = 0.045). This lack of correlation suggests that the temporal attention that subjects give to political content on screen is not related to the way they temporally attend to the smartphone screen in general. Rather, as Subject 46 demonstrates, a smartphone user who distributes their smartphone usage across usage sessions of varied durations may consistently restrict their attention to political content to extremely short windows. Subject 43 shows something of the opposite behavior. Duration of attention to political content on the smartphone screen is more than a manifestation of overall screen usage behavior.

To summarize this subsection: (1) a typical subject's average experience of political content segment duration is more consistent with that of other individuals than with their own intraindividual experience; (2) segments durations are distributed according to a steeper exponential curve than are usage sessions; (3) there is no correlation between the core distributional parameter underlying the power law distributions of segment and session durations intra-individually, suggesting that the rapidity of segments is unrelated to the temporal structure of overall usage. In the next subsection, I examine how exposure to political content is related to the applications onscreen at the time of encounter.

RQ3.1 How do political content segment frequencies vary across application categories?

As was outlined in Section 4, political content has historically been accessed in a news format, and thus researchers have continued using the presence of that format as a proxy for political content. Simultaneously, social media has been implicated as a source of political content on the smartphone. In this section, I examine how political content encounters are distributed across all application categories, including applications categorized as *news* and *social media*. To do so, I first provide context by examine subjects' *overall* allocation of time across application categories, political or not. As described in Section 8, applications in the sample are categorized primarily according to their designations on the *Google Play Store*, with some manual adjustments to accommodate the specifics of the sample rather than developers' commercial categorization goals. Foreground application data was not collected from six subjects due to technical problems with the data collection application. See the application categorization schema in Appendix D for more details.

In the four panels of Figure 12 (*page 73*), I decompose subjects' screen usage and political content encounters across 23 application categories. In all four panels, the intraindividual distribution of screen activity from 109 subjects is shown as a series of 23 boxplots with underlying subject-level data shown in one-dimensional low-opacity scatterplots. In all four panels, five familiar subjects are highlighted in color for clarity and continuity. In Panel A, I show the proportion of screenshots (i.e., overall screen activity) from each application category. For example, the top row of Panel A shows that Subject 31 (in cyan) allocated[17] just over 30% of their screentime to communication applications, and across 109 subjects, the median percentage of screenshots captured from communication applications was 16%. In all four panels, application categories are arranged on the *y*-axis in descending order of subjects' median time allocation, and the *x*-axis shows percentage on a log scale. In Panel B, I show the intra-category percent of screenshots that contain political content. In Panel C, I show the percent of political screenshots that came from each category. In Panel D, I show the percent of political content encounters that came from each category.

[17] Throughout this section, where I use the terms *allocate* or *source*, I do not mean to imply intentionality on part of the subjects.

Panel A makes clear that subjects have highly dissimilar intraindividual distributions of screentime across application categories. Even in the most common application categories, namely *communication*, *browser*, and *social media*, subjects' intraindividual allocations range from 0% to over 50%. Moreover, every application category hosts two or more outliers in terms of intraindividual time allocation, emphasizing the idiosyncratic nature of even the simplest metrics of smartphone usage. The category *gaming* is unique in that the median time allocation to gaming is just 1%, yet one third of the sample allocates up to the plurality or even majority of their time to mobile gaming. The Screenomics application category refers solely to the data collection application used in this study, which was also relied upon for collecting foreground application information.[18] Notably, *news* applications are highly uncommon, such that only outlier subjects allocate any screentime to *news* applications at all. Subject 43, in red, is one such outlier, spending 10% of screentime on *news* applications.

The fourth most prominent application category is survey-taking, labelled here as *cash/survey* to incorporate similar activities such as Mechanical Turk tasks or betting, all of which serve as smartphone-based micro-level income opportunities. Research surveys, much like those that our subjects answered for us, are commonly distributed for social scientists and market researchers through the same convenient standardized system — in this case, the Qualtrics panel aggregation and distribution system. The subjects in this sample are part of an apparent class of Americans who frequently answer surveys and conduct similar gig work on their smartphones

[18] I manually audited 575 randomly sampled screenshots stratified by 109 subjects for accuracy in foreground application detection. 83% of automated foreground application classification was correct, and for most subjects, all sampled screenshots were accurately labeled according to foreground application or on-screen notification. The remaining inaccuracy explains why some political content exposure is ascribed to the *Screenomics* application category in Panels B, C, and D of Figure 12, and why some outliers in Panel A appear to have spent more than 1% of screentime viewing the data collection application. The majority of incorrect foreground application labels were owed to home screen traversal following exiting an application, as the data collection application captures the *opening* of applications, not the closing.

for small incentive payments. This is new and concrete evidence of a methodological bias which may apply to most non-probability samples and shared-use panel systems used in this domain. This broadly includes data from *Bovitz* (e.g., Druckman & Levendusky, 2019), *Ipsos* (e.g., van Erkel & van Aelst), *Knowledge Networks* (e.g., Cacciatore et al., 2018; Prior, 2010), *Qualtrics* (e.g., Thorson et al., 2019; Yang et al., 2019), *Survey Sampling International* (e.g., Molyneux, 2017), *TargetSmart* (e.g., Iyengar, Konitzer, & Tedin, 2018), and *YouGov* (e.g., Guess, Nyhan & Reifler, 2020; Guess, 2021; Peterson et al., 2021). To the extent that scholars of political communication assume that media effects such as *priming* influence how subjects report their political opinions, the temporal cross-contamination occurring within frequent survey-takers' experience on these platforms is worth scientific study in its own right. In my main analysis, I have chosen to include these survey-based encounters in my estimates and analyses of political content screenshots, encounters, and segments, as they represent the genuine and natural experience of the subjects I aim to describe. However, in Appendix I, I show the results of all core findings with the omission of survey-taking by removing survey-taking screenshots from the screenshot record outright. My interpretation of these alternative results relative to results in the main analysis is that they do not significantly alter the substantive findings of my dissertation but have more serious consequences for other analyses common in the field.

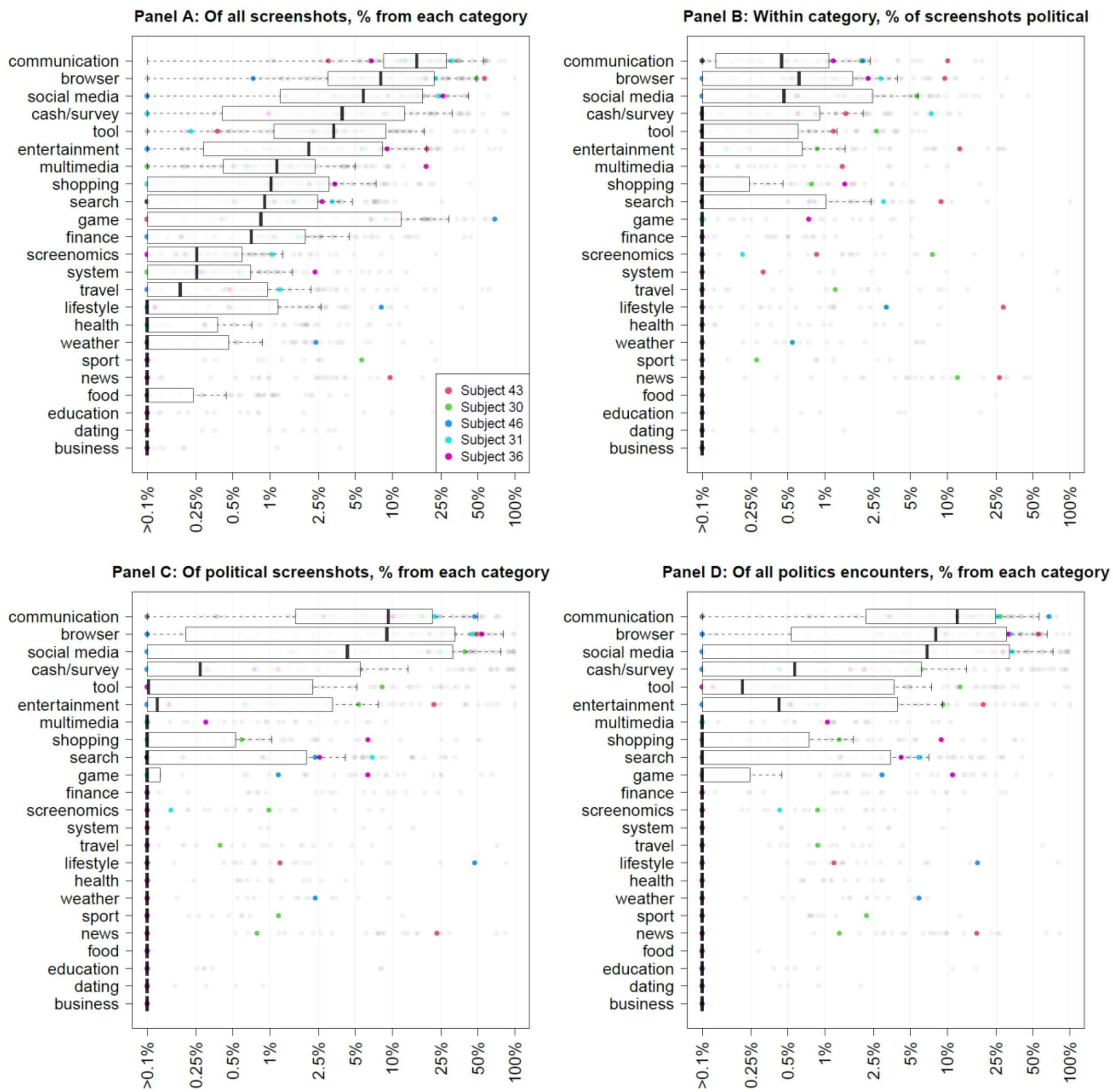


Figure 12. *Decomposition of overall screen exposure and exposure to political content across 23 application categories for 109 subjects with foreground application data. Each subject is represented as a single point 23 times in each panel, with boxplots summarizing subjects' values within each application category. X-axes are explained in panel titles and shown in log scale; y-axis values are 23 categories in descending order according to medians in Panel A. Five subjects are highlighted in color for clarity and continuity with previous figures. Categories are based mainly on Google Play Store labels; see Appendix D.*

Moving on, Panel B shows the percentage of screenshots within each category that contain *political content*, with boxplots again summarizing the underlying subject-level datapoints. *Communication*, *browser*, and *social media* applications are not only the most used applications overall, but also contain the highest percentage of political content in terms of median subject-level time allocation. Even still, median time allocation to political content within any category is under 1%, and for most categories, the median subject spends 0% of screen time exposure to political content. Apart from three outlier subjects, no application category is used mostly for political content consumption by any subject; of those three subjects, the application category in question is not *news*. For all subjects who used *news* applications, the majority of screenshots captured while *news* applications were on screen were not classified as containing political content. Moreover, the frequency of political content exposure within the *news* category is similar to that of application categories not ordinarily associated with political content, such as *entertainment* and *lifestyle*. In the smartphone environment, these categories are all similarly likely to provide access to political content, notwithstanding wide variation across subjects' individual-level behavior. Subject 43, the subject who was exposed to the most political content in the study, spent 27% their *news* application screentime on political content. However, they were also an outlier in several other application categories, implying that political content was a common feature of a wide array of formats and spaces in their personal smartphone environment.

The category *gaming* is unusual in that a large amount of overall screentime is dedicated to it, yet very little political content is garnered from it.[19] Conversely, while comparatively little overall time is spent using *search* applications, a large share *search* application usage involves exposure to political content — though for the median individual this percentage is zero. Altogether, political content appears in every app category at least once in the sample, but across subjects, the median experience of most application categories (including *news*) is devoid of

[19] Outliers in this are presumed to reflect the occasional occurrence of political or warlike terms in games, such as the word *campaign* occurring in the game *Clash of Clans*.

political content. To clarify the importance of this finding: within this sample, if a study were to proxy political content exposure by *news* application usage, such study would be *mainly* measuring non-political activity and actually ignoring most exposure to political content.

In Panel C, I show how each subject allocates their political content exposure across application categories. In this plot, the sum of *x*-axis values is 100% for each subject. As with Panel A and Panel B, Panel C shows that subjects' intraindividual distributions are widely varied. For the median subject, 8% of political content exposure occurred via *communication* applications, another 8% via *browser* applications, 3% via *social media*, and negligible amounts from all other application categories. While *browser* applications may host any type of content available via URL, *communication* applications provide political content via incoming campaign emails, political newsletters, text conversations, memes in group chats, and so on. In stark contrast, the median subject sourced 0% of their political content exposure from *news* applications, though several outliers such as Subject 43 sourced the plurality from *news* applications. In the entire sample, only two subjects sourced the majority of their political content exposure from *news* applications. Some subjects source the majority of their political content exposure from applications that are not traditionally associated with political content, including *cash/survey* applications but also *games* and *tools* wherein political content may arise in gameplay, advertisements, or even notifications. For Subject 43, roughly half of political content exposure comes from *browser* applications, with the other half split between *news* applications and *entertainment*. Lastly, Panel D shows how each subject allocates their political content *encounters* across application categories. This view illuminates the unique instances and opportunities that subjects have to engage with political content, regardless of exposure time. The similarity between Panels C and D in part reflects that 44% of political content encounters do not lead into segments longer than five seconds. The only exception is within the *communication* category, in which the median subject sources 12% of their political content encounters versus 8% of their political content exposure. This difference implies that a sizable share of political content encounters occurring via communication channels do not result in engagement, as proxied by exposure lasting more than one screenshot.

Through various decompositions, Figure 12 shows that application categories used to access political content include a broad range not easily captured in aggregate. As discussed in Section 4, political communication research has long conflated political content (a variety of content) and news (a format); the reasonableness of that conflation has sharply declined in the past decade. To address this concern directly, Table 6 is a cross-subject mean breakdown of screenshots containing political content or not with screenshots containing a *news* application or not. While there is variation across subjects, for the mean subject only 1.13% of screenshots came from news applications, and just 0.15% of screenshots contained political content via news applications. Rather, 1.68% of screenshots contained political content accessed via other application categories, meaning only 9% of exposure to political content occurred while a news application was on-screen, for the average subject. To the extent that *news* applications are the successor of traditional news formats, the study of politics via news in the smartphone environment is focused very narrowly on the 0.15% of screenshots in the average subject in this sample. This finding strongly suggests that *news* applications are an inappropriate tool for measuring exposure to political content on smartphones, due to low measurement validity.

Contingency table: *News* applications versus political content			
	Political	Not political	Total
News	0.15%	0.98%	1.13%
Not news	1.68%	97.19%	98.87%
Total	1.83%	98.17%	100%

Table 6. Contingency table of political content and news application usage. Values shown are the average of intraindividual percentages across 109 subjects with foreground application data. Political content is classified by textual content classification; news formats are classified according to application.

Table 6 uses interindividual averaging to provide a straightforward contingency table. However, there is very strong interindividual variation masked by averaging. In Figure 13, I graphically demonstrate all 109 subject-level contingency tables underlying Table 6. The *y*-axis shows the 109 subjects ordered by overall exposure to political content as a percent of total screentime. The *x*-axis shows percentage of intraindividual screentime. Red represents political content exposure via any application category besides *news*. Yellow represents *news* application usage without political content detected on screen. Orange represents political content exposure occurring on news applications, and gray represents all other screen usage. Five subjects are highlighted in color for clarity and continuity with prior figures.

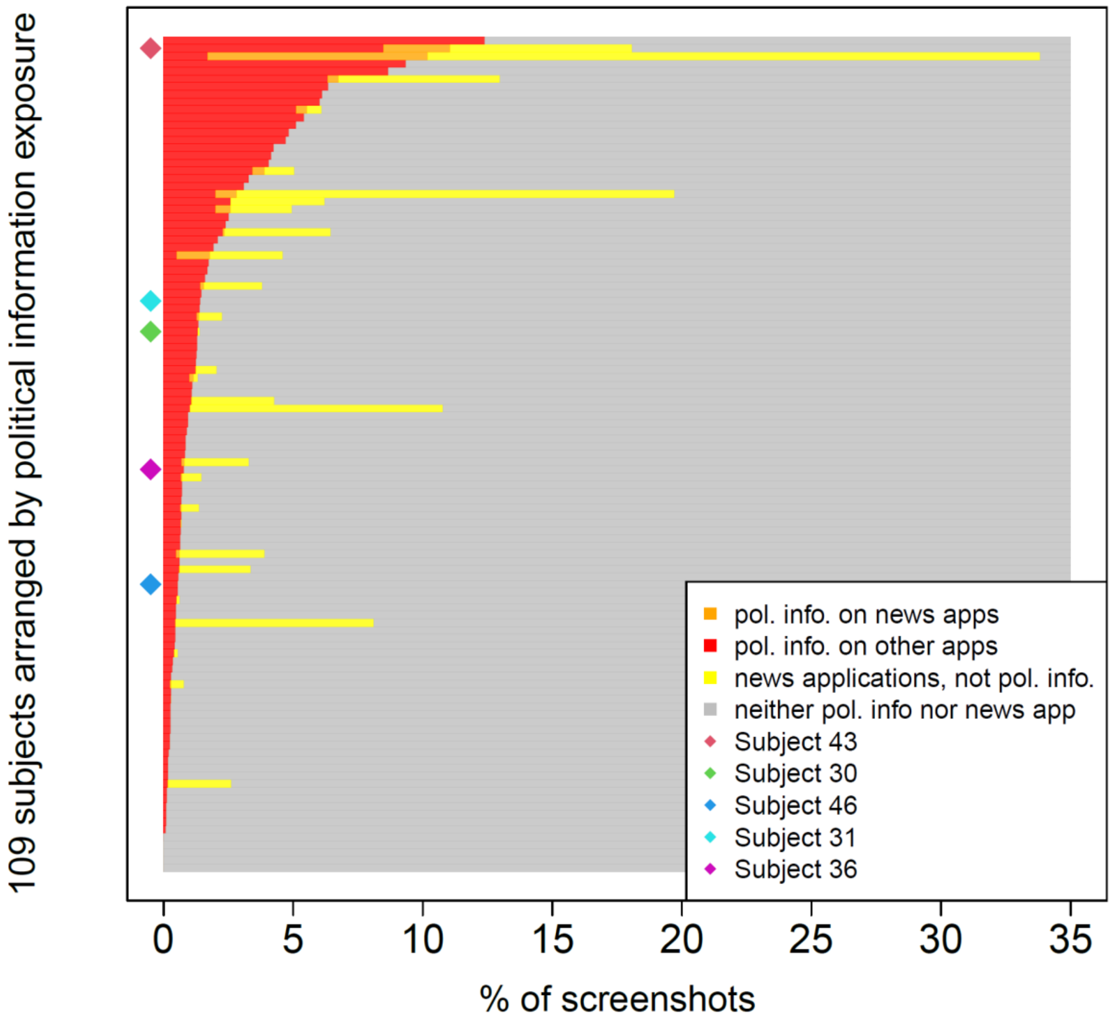


Figure 13. Graphical representation of 109 subjects' contingency table of political content and news within the screenome. Political content is classified by textual content classification; news formats are classified according to application. Each horizontal row in the plot shows a single subject's intraindividual experience across their entire data collection period, with colors labeled in the inset legend. Five subjects are highlighted in color for continuity with prior figures. News applications are used by a minority of subjects; political content is encountered by most; the two rarely overlap.

Figure 13 illuminates several points that cannot be deduced from an aggregate contingency table. Most subjects never open a *news* application. Of those who do, none use them primarily for political content. As was also seen in Figure 12 Panel C, just two subjects source the majority of their political content from news applications. Here we see that these two seemingly similar individuals have mutually distinct news application usage and political content exposure. The interindividual correlation between news application usage and political content exposure is just 0.37; dropping Subject 43 lowers this correlation to 0.32. The validity risk of proxying political content exposure with news applications is not merely a problem of bias but also of strong interindividual variation without clear structure.

To summarize this subsection: (1) on the smartphone, news exposure is a very poor proxy for political content exposure; (2) for most subjects, political content is dispersed across several categories; (3) intraindividual distribution of political content exposure across application categories is extremely varied across subjects; (4) political communication researchers risk cross-contamination of studies by driving a finite set of subjects to complete surveys with political content. In the next subsection, I switch from examining exposure and frequency to *durations of political content* segments with respect to application categories.

RQ3.2 How do political content segment durations vary across application categories?

As already described in this analysis, segment durations are tightly clustered toward the minimum possible duration, five seconds, across all subjects. In general, the prevalence of shorter segments signals political content is often not engaged with when it is encountered on screen. Here, I examine whether durations vary according to application categories. If a category tends to bring forth shorter segments than does another category, this suggests that the former category is a comparatively lower-quality source of political content. This analysis requires analyzing vectors of segment durations across applications, subjects, and application categories simultaneously. To simplify this analysis, I reduce the set of application categories from 23 to five*: news*, *social media*, *communication*, and *browser*, with all others being grouped together as *other*. These categories were chosen because of the relevance of *news* formats to political communication research and revealed preeminence of *social media*, *communication*, and *browser* applications in producing political content exposure.

Figure 7 (page 57) showed each subject's intraindividual distribution of political content segment durations, as well as a mean curve summarizing the experience of all subjects. In Figure 14, I apply this same approach, broken down by five application categories. Axes are shared by all six panels: along the *x*-axis, I show segment durations from 5 seconds to 60 seconds, in increments of 5. Along the *y*-axis, I show the percent of political content segments lasting *X* duration. In the smaller panels along the bottom of Figure 14, I show the intraindividual per-category distributions of political content segment durations, with colors as indicated in the legend in the main panel. In the main panel of Figure 14, for each application category, I summarize all subjects' intraindividual distribution of political content segment durations. I do so by calculating cross-subject mean percentage at each segment duration along the *x*-axis and adding two standard errors above and below, visualizing basic confidence bands for each application category. Ultimately, these confidence bands capture interindividual variation of intraindividual distributions of political content segment distributions across 0 to 60 seconds in 5 application categories.

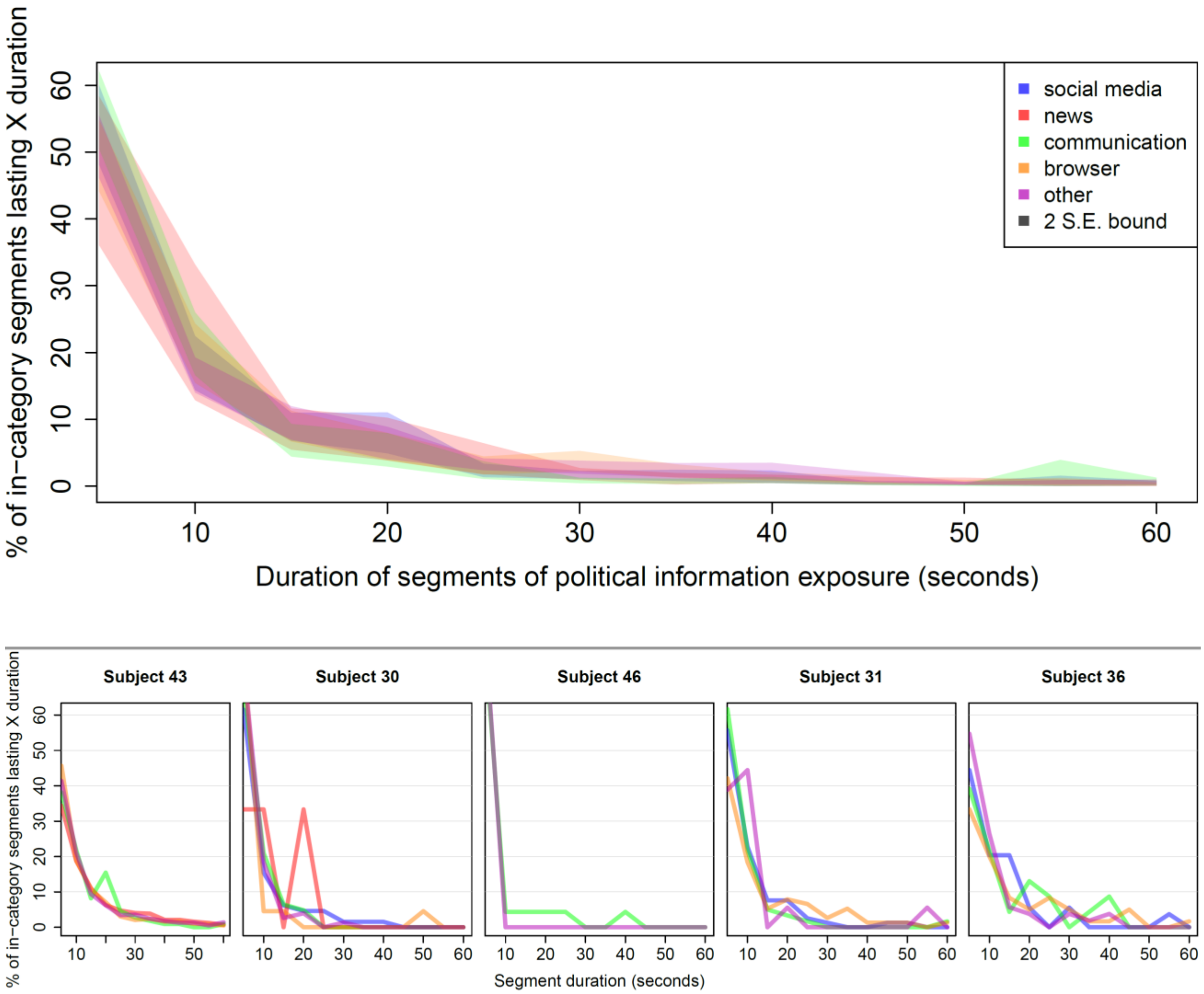


Figure 14. Histograms of segment durations according to application category. In the main panel, for each of five application categories, I show the interindividual mean of intraindividual histograms of segment durations, from 5 seconds to 60 seconds along the x-axis, +/- two standard errors pointwise to create full confidence bands in low opacity. Pairwise comparisons between the five application categories at each duration level find that intraindividual segment durations are not significantly different in virtually all cases (Appendix E); this is reflected in the overlap between all five bands across the entire plot. Five familiar subjects' intraindividual category-specific histograms of political content segment durations are shown in smaller panels for clarity and continuity with prior figures. All colors are labeled per the inset legendin the main panel. Percentages reflect all segment durations, including those beyond 60 seconds long. This plot is zoomed into the origin, visually omitting a very long rightward tail.

The main panel of Figure 14 shows that intraindividual distributions of political content segment durations are very similar across application categories. This is demonstrated visually by the near-complete overlap between all five colored bands. Appendix E provides a full pairwise statistical comparison of each category pair at each duration level; virtually all such comparisons are non-significant. Of all application categories, segments from *news* applications show the widest variation in intraindividual duration distribution. Still, for the mean subject, nearly half of political content segments occurring within a *news* application last just five seconds. For all other application categories, this value is just over 50%, with *communication* being the most likely category to produce five-second-long segments rather than longer segments, for the mean subject. These intra-category estimates are similar to the whole-sample estimate of Figure 7 (page 57).

Figure 14's main panel illuminates a crucial point about political content exposure on the smartphone. Across subjects, political content segment duration is not conditional on application category. Though each subject makes use of a unique repertoire of applications and experiences a unique degree of overall exposure to political content, segment durations are consistently distributed according to a steep power law curve focused at the shortest duration. The small panels along the bottom of Figure 14 help clarify this. Subject 43, on the left, encounters a large array of political content on *browser* applications, *news* applications, and *communication* applications. In all three categories, their political content segment durations are distributed along a well-defined curve, with the plurality of segments lasting just five seconds. The remaining four subjects are exposed to less political content (in descending order: Subject 36, 31, 30, and 46), so their intraindividual segment duration distributions are less resolved. However, apart from limited resolution, all curves follow a similar curve unconditional on application category.

Segments lasting longer than 60 seconds are rare, hence their omission from Figure 14. However, longer-lasting segments suggest greater likelihood of engagement and information retention, as they imply recall rates greater than 50% (see Table 1, page 27). Across the sample, there were 752 segments lasting longer than one minute, or 2.9% of segments. These longer segments were experienced by 75 out of 109 subjects with foreground application data and occurred in 19 of 23 application categories. This wide array of subjects and application categories means that long segments are not simply driven by few avid subjects or by a few especially conducive applications. In Figure 15 I examine this point further. For each of eight application categories (one of which is a catch-all "other" category), I plot the median duration of long segments (along the *x*-axis) against the number of long segments (along the *y*-axis). Points are sized proportionally to the number of unique subjects who experienced a long segment within the application category. The three additional application categories added relative to the prior figure (Figure 14, page 81) are *entertainment*, *survey/cash*, and *search*, each of which had a comparatively large share of long segments relative to other application categories.

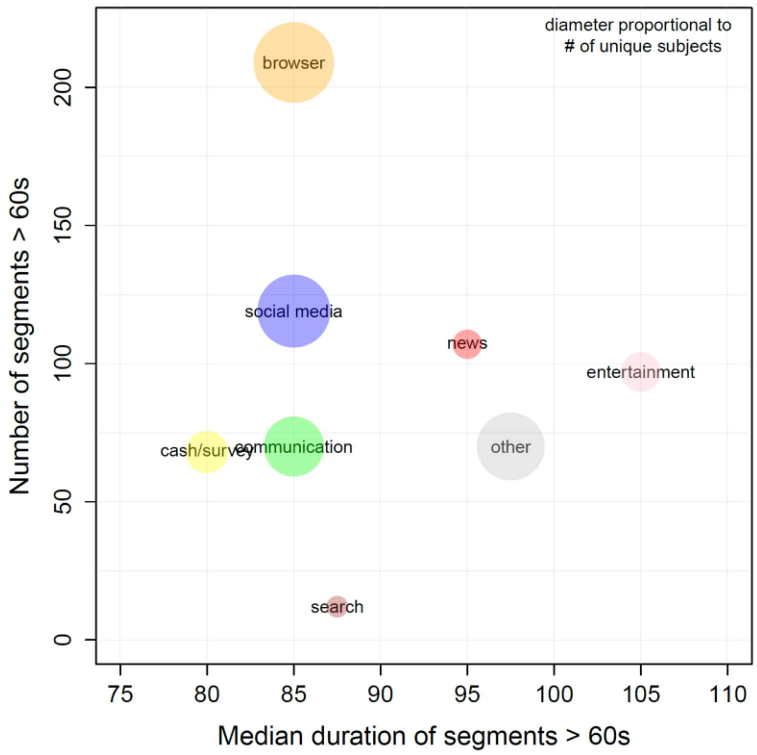


Figure 15. Long segments (segments of political content lasting longer than 60 seconds) grouped by application category. The x-axis indicated median intra-category duration of long segments, in seconds. The y-axis indicates the count of intra-category long segments. Point size is proportional to the number of unique subjects experiencing a long segment while using applications in the category.

Figure 15 shows foremost that the median duration of the 752 longest segments is fairly short. Even when considering only segments lasting longer than one minute, the median duration of segments from any application category is under two minutes, in line with the power law shape which segment distributions have been shown to follow. Browser applications are the most prolific source of long segments, both in terms of overall segment tally and number of unique subjects. Long segments on entertainment applications including *YouTube* have the longest median duration yet occur for a comparatively small set of subjects and in comparatively fewer instances. Long segments occurring via *news* applications occur for fewer unique subjects than those occurring via nearly all other applications but are more common in the sample than long segments by any category besides social media and browsers. This speaks to the narrow importance of news applications to a small subset of users, as was visually demonstrated in Figure 13. The median duration of long segments occurring via *cash/survey* applications is just under a minute and half, matching expectations about the duration of online surveys. These long segments occur for more unique subjects than do long segments via *news* applications despite occurring less frequently in the sample. Overall, long segments are very rare, are spread across multiple subjects and multiple application categories, but are more common via the browser than any other application category.

Lastly, I examine intra-segment cross-category traversal, which could signal engagement even at shorter durations. A political content segment that is initiated by a political content encounter can occasionally traverse many applications. As shown in the Figure 2 diagram, political encounters on social media can feasibly lead into a segment mostly contained in another app category. Cross-app traversal, however, is extremely rare. Out of 26,090 segments for which foreground application data is available, just 443 (0.2%) traverse multiple application categories. The median duration of these 443 segments is 30 seconds, versus ten seconds in the entire sample. Out of 443 segments, 404 traverse two application categories, with remaining segments traversing three, four, or (in two cases) five unique categories. A list of all cross-category traversal segments, and their durations, is provided in Appendix F.

To better understand the unusual phenomenon of cross-category segment traversal, I present a bipartite network graph in Figure 16 representing all 404 segments that began in one application category and ended in a second application category. In the left column of nodes, I list the starting application category, and in the right column of nodes, I list the finishing application category. Edges are colored according to the starting category and weighted according to the number of segments following the traversal flow.

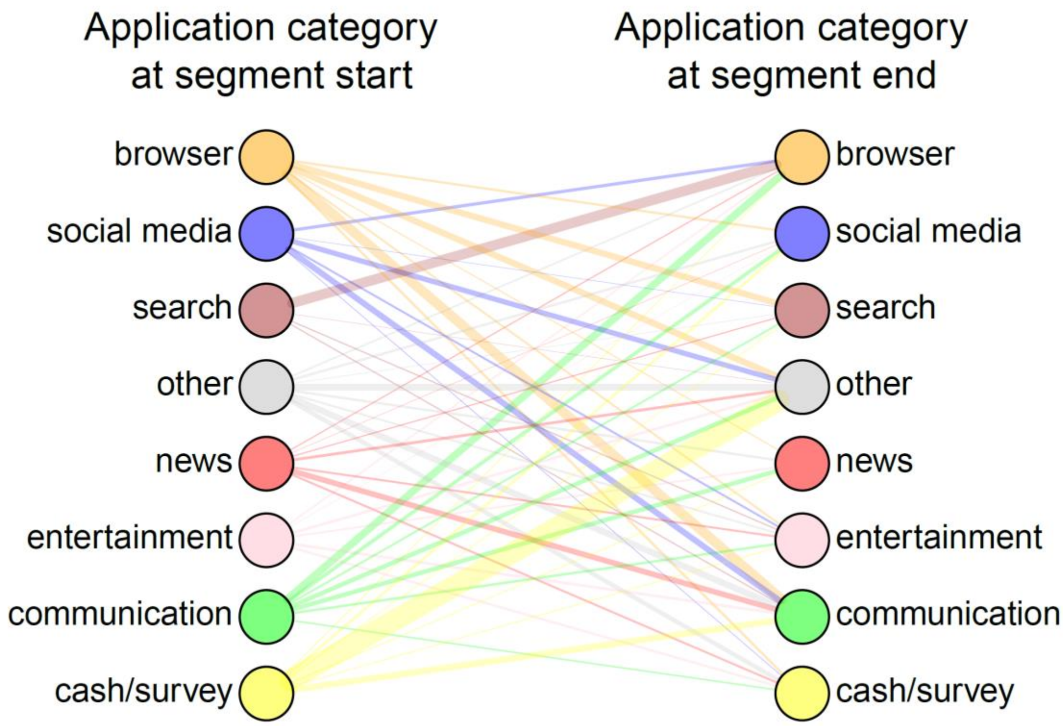


Figure 16. Directed bipartite network graph showing intra-segment traversal across application categories for 404 political content segments (out of 443 total known to have traversed more than one segment; 29 traversed >2 categories). Edges are directed left to right and colored according to the application category in which the segment began. Edge width is proportional to the number of segments which follow the shown traversal.

Figure 16 shows that there is no single source or destination for cross-category segment traversals. Cross-category traversals include all seven application categories in which political content segments tend to occur more generally. Still, there are a few intuitive results visualized within the graph. While the *search* application category is not among the top four source categories for political content segments, it created 33 cross-category traversals, 26 of which went to *browser* applications. These 26 segments represent the most concretely intentional instances of political content exposure in the entire sample, given that *search* applications are used for locating information on the *browser*. Sixty-five segments traversing from *communication* applications to other application categories suggest that subjects followed inbound links shared via email inboxes or messaging. Segments ending in *communication* applications (96) or *social media* applications (31) suggest the opposite: political content being shared with others. Ultimately, traversals are extremely rare in the sample, reflecting the already-low frequency of political content exposure, and they are not driven by any singular application category.

To summarize this subsection: (1) intra-individual distributions of segment durations are not conditional on application category; (2) long segments (operationalized as segments lasting more than one minute) are rare and driven mainly by *social media* rather than *news* applications; (3) only 0.2% of segments traverse multiple application categories, with no single application category being a primary source. In the next subsection, I switch from examining the role of application categories to examining the role of text features and their sequencing in political content exposure.

RQ4 How do segment durations vary according to valence, arousal, and wordcount preceding and during segments?

My final research question regards sentiment and textual content of political content segments. As discussed in Section 5, arousal and valence are known to influence how media users attend to media. In the context of smartphone usage, content has been shown to follow a sequencing pattern of self-regulation, with higher arousal content preceding lower arousal content (and vice versa), and more-informational content preceding less-informational content (and vice versa) (Cho et al., 2022). In other contexts, counter-normative content has been shown to elicit stronger responses and reactions from media users (Schwager & Rothermund, 2013). Still, the role of arousal, valence, and informational content has never been studied in political content exposure on the smartphone, especially in context of temporality. Should content sequencing play a role in segment duration, this implies that the variations in information processing associated with variations in segment duration are in turn related to rapid and inherently idiosyncratic regulatory processes.

To begin, I calculate the wordcount, arousal, and valence level of every screenshot in the sample. This was done using the National Research Council of Canada (NRC) valence, arousal, and dominance dictionary, an open source 20,000-word emotive lexicon based on Mechanical Turk ground truth generation (Mohammad, 2018). Valence and arousal levels in this lexicon range from 0 to 1. An arousal score of 0 indicates that a word excites the least possible degree of arousal, while an arousal score of 1 indicates a maximally exciting term. A valence score of 0 indicates that a word conveys maximally negative valence, while a valence score of 1 indicates maximally positive valence.[20] For each screenshot, all words on screen were isolated (by single spaces) and merged with the lexicon. Valence and arousal levels were then averaged within-screenshot. Arousal and valence scores for words not found in the lexicon were treated as missing values. For each subject, screenshot-level valence and arousal scores were standardized such that each subject's mean screenshot valence and arousal value is set to zero, with standard deviation being one. The same intraindividual standardization was done on screenshot-level wordcount. This allows for comparison of behavior across subjects and segments while controlling for subject-level biases. Standardization also enables simplified visualization on shared axes, as in Figure 17 below. Wordcount, valence, and arousal values for segments are calculated as the mean of screenshot-level values within the segment.

[20] Others using the circumplex model have treated valence as two independent variables, positive (0 to 1) and negative (0 to 1) with ambivalence being possible. In the NRC model, these positive and negative scales are considered one and the same, attached at the origins where valence = 0.5 on the lexicon's default scale.

In Figure 17, axes are shared across all three panels. The *x*-axes represent segment durations lasting between five and 120 seconds; the *y*-axes represent intraindividual standard deviation from the intraindividual mean value for three variables using a shared standardized scale: wordcount, valence, and arousal. In all three panels, bands represent pointwise mean values of all segments of duration *x* in the sample, +/- two pointwise standard errors. Overlap between colored bands does not convey meaning, as bands represent three different variables. Overlap with the horizontal zero line indicates no significant difference from intraindividual overall mean value of each variable. The increased width of bands toward the right of each panel reflects the exponential drop in number of longer-lasting segments, resulting is much larger standard errors. Descriptive ordinary least square regression lines are overlaid in corresponding colors; coefficient values & corresponding significance levels are shown in-panel, using clustered standard errors by subject. Note that segment durations are measured as a whole, such that an *increase* in segment duration refers to a segment duration of larger size, rather than a marginal increase on an ongoing segment.

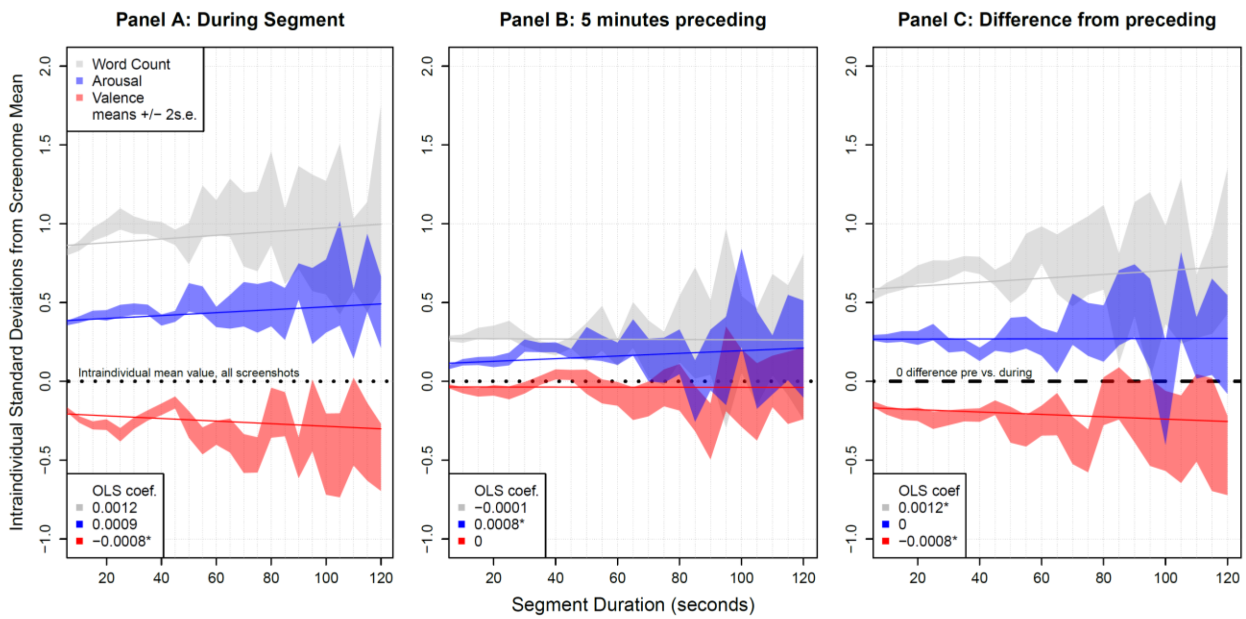


Figure 17. Segment duration versus wordcount, arousal, and valence during and preceding the segment. X-axes represent segment durations lasting between five and 120 seconds; y-axes represent intraindividual standard deviation from the intraindividual mean value. Bands represent pointwise mean values of all segments of duration x in the sample, +/- two pointwise standard errors. Overlap with the horizontal zero line indicates no significant difference from intra-individual mean value of each variable. Descriptive ordinary least square regression lines are overlaid in corresponding colors, with coefficient values and their significance level shown in-panel (: p < 0.5, errors clustered by subject). Valence and arousal are based on NRC VAD lexicon (Mohammad, 2018). Segment durations describe whole segments, not marginal increases. See Appendix G for extended analysis.*

In Panel A, I plot the average wordcount, arousal, and valence of political content segments across durations. Starting at the left side of the panel, Panel A shows that political content segments lasting five seconds have significantly higher wordcount, significantly higher arousal, and significantly more negative valence than the average screenshot, standardized per subject. Each of these results is sensible. First, regarding heightened wordcount, text format is a common method of presenting political content. Recall that just 2.4% of political screenshots in the ground truth set contained fewer than five text characters, and in turn, my political content classification strategy relies on the presence of text on screen. Regarding heightened arousal, this is interpreted to reflect the intensity of political topics relative to other topics. Regarding negative valence, this is interpreted to reflect the gravity of political topics relative to other topics (Quattrone & Tversky, 1988; Soroka, 2006).

Ordinary least squares regression shows that there is a slight but significant negative relationship between segment duration and valence within political content segments. As indicated in the bottom left corner of Panel A, a one-second increase in political content segment duration corresponds with a valence decrease equal to -0.0008 intraindividual standard deviations. There are non-significant positive relationships of similar magnitude between duration and arousal and wordcount. Interestingly, the relationships between segment duration and each of these three variables are *amplifications* relative to intraindividual screenome mean. For example, average wordcount is higher for screenshots within political content segments than for screenshots not within political content segments, and this is gap widens for increasingly long political content segments. Regarding valence, the amplificative relationship corresponds with findings from Hsu et al. (2022) which suggest that American media users have an attentional bias toward counter-normative, i.e., negatively-valenced content. More generally, cognitive involvement with short on-screen content (microblogs) has been shown to be associated higher arousal and more intense valence (Stieglitz, & Dang-Xuan, 2013). Such a relationship, however slight, appears to exist in the case of political content segments on the smartphone, though the causal relationship between duration and sentiment is unknown.

Figure 17 Panels B and C illuminate how segments may be reactive to preceding content, in line with Cho et al. (2022). First, in Panel B, I show the relationship between segment duration and the average wordcount, valence, and arousal, respectively, of the five-minute period immediately preceding political content segments. The choice of five minutes is arbitrary, but results are robust to alternative preceding durations; see Appendix G for an extended analysis.

For segments up to seventy seconds long, wordcount and arousal during the preceding five minutes is significantly higher from the screenome mean. These values follow the same pattern as wordcount and arousal within segments, but to a lesser extent. Overall, Panel B suggests that most political content segments arise from usage periods in which content is structured similarly to political content: informative, with high arousal, and with valence that tends toward negativity, though very slightly. More interestingly, this is not true of longer segments. In line with the concept of self-regulation, political content segments tend to last longer than one minute if they occur following a five-minute period of screen behavior *not* significantly different from the screenome mean. Still, a stricter interpretation of self-regulation would suggest that the minutes preceding a long political content segment would exhibit patterns *inverse* to those of political content: lower-than-average wordcount, lower-than-average arousal, and positive valence. Such a relationship is not found. Regression coefficients linking segment duration to wordcount and valence are insignificant. There is, however, a significant positive relationship between arousal level in the preceding five minutes and the duration of a segment. The more exciting the *on-screen text preceding a segment of political content* is, the longer the political content segment is expected to last.

Figure 17 Panel C more directly examines the contrast between segments and their preceding five minutes by taking the difference between the two. In Panel C, the horizontal zero line indicates no significant difference between intra-segment and pre-segment values of wordcount, valence, and arousal respectively. essentially recreates the patterns of Panel A. That is, political content segments have higher wordcount, higher arousal, and more negative valence than their preceding five minutes, just as they do in comparison to the screenome mean. The regressions in Panel C suggest strongly that sequencing patterns play some role in attention to political content. First, there is a significant positive relationship between segment duration and *higher wordcount in the segment relative to the preceding five minutes*. That is, the bigger the jump in wordiness when entering a political content segment, the longer the political segment is expected to last, correlationally. Second, there is a significant negative relationship between segment duration and *more negative valence in the segment relative to the preceding five minutes*. That is, the bigger the drop in valence[21] when entering a political content segment, the longer the political segment is expected to last, correlationally. While these is no such significant relationship found regarding arousal, the alternation patterns detected here cohere with fast and apparently unconscious self-regulation found by Cho et al., (2022).

Moreover, Panel C suggests that segment duration is extremely sensitive to these sequencing patterns, notwithstanding causal direction. For the average subject, a jump in average wordcount of just one-tenth of a standard-deviation relative to preceding content corresponds with an 83-second increase in a segment duration; a drop in average valence of just one-tenth of a standard-deviation relative to preceding content corresponds with an 125-second increase in a segment duration. This result has fascinating implications for the study of political communication in the smartphone environment. Since a segment's duration guides the timescale of processes that can occur in reaction to the segment's political content, especially leading to information retention, then the degree of political learning that occurs via the smartphone in each instance is apparently related to the rapid and idiosyncratic media sequencing within which the segment is embedded.

To summarize this subsection: (1) political content segments have significantly higher wordcount, higher arousal, and more negative valence than the intra-subject screenome mean; (2)

[21] Again, valence is measured as a single spectrum from intensely negative to intensely positive, in line with NRC structure.

these differences widen significantly at longer segment durations; (3) political segment length correlates positively with greater contrast in wordcount and valence relative to the five minutes preceding the segment; (4) segment duration is extremely sensitive to text features of preceding content, correlationally; (5) variation in information processing implied from variation in segment duration is transitively related to rapid and idiosyncratic sequencing of content. The next and final section is a concluding discussion.

Section 10. Discussion

In this section I summarize my findings, discuss their implications with respect to my dissertation goals, and explain their relevance to assumptions listed in my introduction. I also discuss research limitations and directions for future research.

Findings Summary

Through a range of research questions and analyses, my dissertation has aimed to provide a missing description of political content exposure in the smartphone environment. Here I provide a high-level summary of my findings. Implications, limitations, and future research follow.

Political content appears infrequently in the sample overall. Out of 4,907,091 screenshots captured across all 115 subjects, 92,988 contained political content. This is equivalent to 1.89% of screenshots. The number of political content encounters (and equivalently, the number of segments) was 26,238 spread across 12,753 unique sessions. Most exposure to political content is driven by a small share of subjects, suggesting that group-level means aiming to describe a single typical individual's experience fail to capture the exponential structure of interindividual exposure levels, and thus describe essentially non-existent people. The highest percentage of screentime in which political content was on screen was 11% for the top subject, while the median percentage was 2%. Subjects' intra-individual variation in political content exposure, measured as unique encounters or as percent of screentime, is on par with inter-individual variation. On average, a subject's daily political content exposure is inaccurate by a factor of 4 in estimating any particular hour of the subject's experience, in expectation. Intraindividual estimates of political content exposure are increasingly inaccurate when applied at shorter time spans, tracing the value of time scale explication in measurement.

Nearly half of segments last only the shortest possible detectable duration, five seconds. The expected recall and recognition rate of the content of most segments is under 15%, based on advertising research. Only a negligible amount of political content segments last longer than one minute. Considering the power-law curve of the segment distributions revealed throughout the analysis, there is good reason to speculate that a very large number of segments last even shorter durations but are missed by detection granularity. Segment durations are distributed according to a steeper exponential curve than are usage sessions. However, the rapidity of segments is unrelated to the temporal structure of overall usage, based on comparisons of distribution parameters per subject. Moreover, a subject's inter-hour average political content segment duration distribution is more consistent with that of other individuals' hourly averages than with their own intraindividual experience in each hour, in expectation. Bivariate relationships between exposure duration, encounter frequency, and average segment duration are ergodic at the hourly and daily levels.

Exposure to political content is not limited to *news* applications. Rather, just 0.15% of an average subject's screenshots contain political content withing a *news* application. Intra-individual distributions of segment durations are not conditional on application category. Long segments, operationalized as segments lasting more than one minute, are extremely rare and are driven mainly by *social media* rather than *news* applications. Only 0.2% of segments traverse multiple application categories. There is high variation across subjects in terms of allocation of political content exposure across application categories. Repertoires are different for each subject, but focus on *social media*, *communication*, *browser*, and *survey-taking* applications. This last application category challenges the use of convenience samples: political communication researchers risk cross-contamination of one another's studies by driving a finite set of subjects to complete multiple surveys with political content.

Political content segments have significantly higher wordcount, higher arousal, and more negative valence than the average screenshot, standardized by subject. These differences widen significantly at longer segment duration, though causal direction is not clear. Segment length correlates positively with a jump in *wordcount* from the five minutes preceding the segment, and with a drop in valence. In these relationships, segment duration is extremely sensitive to variation in the text preceding and during segments, correlationally, suggesting that platform-enabled rapid sequencing plays a role in how political content is engaged with.

Implications

In this subsection, I place my findings in context of implicit assumptions made in political communication literature, drawing on my exploration of prior literature in Section 3, Section 4, and Section 5.

First, the conventional approach of treating *unique instances of exposure* as commensurable units of analysis is not appropriate for the study of modern media and is rather a holdover from prior eras of mass communication (Lazer, 2020; Tewksbury et al., 2001). Barring very particular research agendas, methods that treat encounters as units make a strong or weak assumption about the encounters. The stronger assumption is that each encounter leads to a similarly meaningful segment of exposure, where ‘similarly meaningful’, however construed for a particular research goal, largely implies durations of similar magnitude. My results suggest that this assumption is untenable. I find that 44% of all political content encounters on the smartphone lead to segments lasting just five seconds long, with the possibility for far more encounters of even shorter duration captured between five-second intervals. Advertising research in the most similar media environment suggests that the content of these encounters is hardly retained, relative to segments that last longer. From this minimal duration of five seconds, I find that political content segment durations follow an exponential-shaped distribution such that 61% of encounters lead to just 17% of all time spent exposed to political content on the smartphone. Average screenshot-level wordcount is significantly positively associated with segment duration, suggesting that longer segments pack more information and may thus be more valuable in terms of engagement with political content. Combined with known effects of exposure duration on retention, this implies that even purely intraindividual tallying of political content encounters leads to a misrepresentation of subjects’ genuine experience.

The alternative weaker assumption is that the duration of political content segments is randomly distributed across and within subjects, meaning that estimates of exposure based on encounters are nonbiased in expectation. My results suggest that this assumption is more appropriate, but not definitively true. Distributions of political content segment durations are very similar interindividually, and intraindividually they are not conditional on application category. However, there may be other structural features underlying the occurrence of longer or shorter political content segments that are directly related to research goals. For example, researchers aiming to understand partisan information consumption must assume that segment durations are distributed equally between the two types of content — an assumption that seems somewhat unreasonable. The significant positive association between segment duration and *arousal in the five minutes preceding* a segment makes clear that segment duration is associated with features of smartphone usage outside of political content, meriting further investigation.

A second assumption common in political communication research is that segment durations sum linearly to a measure of total exposure either intraindividually or interindividually (Allen et al., 2020; Guess, 2021; Muise et al., 2022a). Measurement strategies which aggregate political content exposure without regard for the duration of individual segments risk equating entirely different psychological processes per the experience of the smartphone user. (Lemke, 2000). In this case, a subject's total exposure time to a particular source of information (e.g., news consumption in a month, or exposure time on one partisan source versus another) is assumed to be a straightforward additive composite of all individual segments. This approach is more rigorous than an encounter-only approach in terms of measuring information dosage. However, given the extremely short durations of most segments in smartphone environment, and the low feasibility of information retention and engagement from these segments, linear summation still poses a bias risk in any exposure estimation. Evidence is not in favor of time units adding linearly toward political outcomes: retention is shown to empirical grow non-linearly with increased exposure, and more theoretically, different processes simply occur at different timescales.

Leniently supposing that recall and recognition rates following a segment lasting ten seconds (approximately 14%) is sufficient for a political communication researcher's outcome measures, then 9.5% of all exposure in this sample is below a threshold for inclusion in linear summation, as they are comprised of five second segments. Supposing slightly less leniently that the necessary recall and recognition rate is greater 14%, then all ten-second segments are invalid, meaning that 17% of all exposure time in this sample is not valid for analysis. Overall estimates of exposure would thus be biased upward, and to whatever extent five-second segments are directly related to outcome variables in a given analysis, there is clear risk of imbalanced bias in estimating exposure to one form of political content versus another (i.e., in estimating exposure to partisan biased content). Researchers engaging in linear summation of political content exposure must consider their own threshold for how long is long *enough* in determining which segments in include or exclude from total counts. This decision is specific to individual research goals.

Third, no single application category is a majority source for political content. This is especially true regarding news applications and is also true regarding social media. As is made clear by Figure 13, news applications do not proxy political content exposure. Even if speculating that all *browser* application activity in this sample is dedicated to navigating news websites (an assumption that is unrealistic), political content exposure is still overwhelming occurring through other sources. Interindividually, most subjects who were exposed to political content were never detected to be opening or looking at a news application. Inversely, the majority of news application usage produced screenshots not classified as containing political content. Across subjects and in aggregate, social media is a much more prominent source of political content than news applications. Still, measuring political content exposure on smartphones by strictly tracking social media behavior would result in an underestimate of political content encounters by 95% for the median subject (see Figure 12). The large degree of interindividual variation in application category repertoires means that any singular application-based approach to measuring whole exposure to political content would be biased in a manner idiosyncratic to each subject.

Lastly, a large share of political content exposure stems from survey taking, smartphone-based gig work, or other cash-incentive platforms. In the sample, there were nearly 200 such applications found. The median subject spent about three percent of their time on these applications, with wide variation and a long rightward interindividual tail. The median subject received 0.25% of their political content from these applications (relative to a median of 0% for news applications), also with a long rightward tail. For several subjects, the majority of political content encounters or political exposure time came from these applications. This doesn't include the possible use of mobile browsers to complete surveys, nor the presence of survey solicitations in their email inboxes or in notifications, all of which would be classified as coming from other application categories, and all of which have been spotted by the author in manual screenome viewing. This finding validates many researchers' intuited concerns about the quality of samples drawn from convenience sampling systems. For political communication scholars, there are obvious implications for measuring crucial outcome variables including political knowledge, political identity, exposure to politics, and so on. Generalizations of these outcomes are clearly biased by repeated exposure to political prompts by researchers. Moreover, media effects researchers who use convenience sample data without acknowledging the influence of repeated survey-taking are paradoxically assuming that media effects are not important factors in their work. With this dissertation, I assert that the risks are now clear, and researchers must proceed with caution when choosing to work with such data collection systems.

Throughout my analysis, I have focused on the wide interindividual and intraindividual variation in virtually all features of political content exposure. While I have shown that three core bivariate relationships are ergodic, I have also demonstrated that the coefficient of variation of political content exposure and average segment duration is similar across and within subjects. All application of intraindividual subject level means (such as daily exposure, as was used throughout my analysis) to individual occasions represents a possible slip into the ecological fallacy along the temporal axis, as much as is true in the cross-sectional domain. Many popular research topics in political communication, chief among them polarization, are long-lasting sequential processes centered on evolution of intraindividual experience. An echo chamber relies on the repeated exclusion of alternative viewpoints over time. Averaging intra-subject experience across time can grossly misrepresent intraindividual variation, especially when frequent sub-interval fluctuation (e.g., in the bias of consumed information) is held as an indicator of healthy information diets and a signal of functioning democratic discourse.

Limitations

In addressing the goals of my dissertation, my analysis uncovered some limitations which offer promising future directions in methodology innovation. First, the fairly small cross-sectional scale of the sample limits my ability to conduct cross-sectional analysis, in-depth demographic segmentation, and clear generalization of the estimates uncovered here. I am also limited in that my sample was collected via convenience sampling, which suffers from biases discussed in the preceding subsection. My dissertation is not the first analysis in our field to use convenience sampling for recruiting; rather it is the first analysis that has concretely exposed the underlying sampling problem. To whatever extent my findings are biased by my sampling strategy, so too are many others in the literature. What is clear is that subjects in my sample are heavy smartphone users with frequent exposure to micro-incentive gig work on their smartphones, some of which is done via surveying on political opinions and political knowledge.

In my sample, subjects were not only very familiar with survey taking behavior, but also were only *Android* operating system users willing to join a longitudinal screen-recording study for $30. How this selection is related to smartphone usage habits is not clear. The in-situ behavioral influence introduced by knowledge of screen-recording is essentially impossible to measure or quantify in this study. The content of interest in this study, exposure to political content, is not assumed to be sensitive enough to each volunteering user that they would endeavor to hide or promote such usage for the sake of an observer, apart from extreme cases (i.e., extremism or hate speech). If subjects were self-censoring non-political behavior, then this would only meaningfully impact estimates of political content exposure as a percentage of overall screentime. I also do not suspect that the omission of iPhone users from the study would substantially affect my findings, as the human experience of these operating systems are ultimately very similar, as are the distribution of their users.

Screenshot data collection can be improved through better collection technology. The application used in this research was created through the collaborative efforts and hard work of various professors, graduate students, and developers designing and coding a data capture system according to the original goals of the Stanford Screenomics Lab (Reeves et al., 2019). Due to this bootstrapped origin (relative to commercial applications which are developed by dedicated teams), the data collection application faced minor but surprising technical errors that interrupted subjects or caused them to attrite from the study. For example, during onboarding, subjects had to swipe left-to-right through three pages within the application, each of which contained text instructions about how to grant permissions to the data collection application. If subjects with lower-resolution screens had set their default font size too large, they were unable to swipe through these onboarding screens, as the larger fonts would push the invisible swiping detection area off the bottom of the user's smartphone screen. In lower-end phones, back-end memory-management applications would halt the data collection application while it was performing as intended. These hurdles and other similar obstacles were (reasonably) not anticipated by researchers and subjects alike. When technical problems like this occurred, subjects either had to notify me via email if they were aware of a problem, or have their technical problem be detected by my periodic manual checks of all subjects' upload behavior. Once detected, errors were fixed through asynchronous email conversations, many of which ended without resolution or otherwise lasted for more than a day.

Separately, the data collection application also provided a stream of data providing information on foreground application changes. As the application collected reliable information only on foreground applications *opening*, and not *closing*, subjects who hovered on their home screens or lock screens after using an application could produce data incorrectly labeled by the detection process. In some cases on-screen notifications were also registered by the data collection application as foreground application changes. This additional detection is reasonable for representing subject's visual attention, but it was not the intended structure of the application. As was explained in my analysis, 6 out of 115 subjects in this sample did not contribute any foreground application data, suggesting a fault with the data collection application on their smartphone models. All this being said, the data collection system used in this dissertation is the culmination of an admirable feat of research engineering and system management. With technical improvement, future sample collections could be even more impressive.

This dissertation did not explore the role of demographic variation in the sample. Initially, this was motivated by the low utility and generalizability of individual demographic characteristics within the small sample. However, this omission was then concretized by a lack of robust survey data, owed to operational difficulty in establishing a bridge between the recruiting system and data collection system used in this dissertation. A full explanation of this technical limitation is provided in Appendix C. Finally, the extreme granular and visual resolution of the data collected for this dissertation pushes social science into unfamiliar territory. Though my analysis calls for greater care by researchers in measuring subjects' behavior exactly as it occurs, the use of screen-recording raises important questions regarding the value of data and the centrality of subjects' data privacy to the work of scientists. Future innovation with this methodology could find a balance between measurement validity for core concepts and privacy preservation for all other variables.

Future research

The findings of my dissertation analysis motivate revisitation of top-cited studies in the field. Many research efforts using digital trace data were created or published with the foresight for future replications and analyses. In cases where original data is available and contains higher temporal resolution than what was utilized in original research efforts, whole analyses can be conducted once again with clear and accurate addressing of the assumptions discussed in this dissertation. In the context of polarization research, a broad array of findings relies on tallies and balances of intraindividual or interindividual partisan-biased media consumption.

Reinvestigation may reveal that many estimates are biased by unaccounted-for intraindividual variation and the presence of rapid segments. This would potentially shed light on the currently ambiguous state of the literature, wherein the existence, severity, and nature of online echo chambers and filter bubbles are not uniformly estimated across research strategies.

Future research could expand the Screenomics approach to other settings. Browser tracking is now commonplace on desktops and laptops, but whole screen recording has not been used for political communication research. Several applications beyond browsers connect desktop and laptop users to the Internet (e.g., email applications, messaging applications, news applications) and the findings presented in my dissertation suggest that these may all be common sources of political content. As more mediated communication occurs in virtual reality and augmented reality in the coming years, the screen recording approach may be able to capture an increasingly accurate portrait of media users' incoming information.

In Section 8, I explained that a political content segment detected in only a single screenshot lasts for five seconds, as this is the mean value of a uniform distribution of segment duration between zero seconds and ten seconds. The exponential distribution of segment durations found in my analysis strongly suggests that there is *not* a uniform distribution of segment durations lasting between zero and ten seconds, or more importantly, between zero and five seconds. The typical screen refresh rate of modern screens is 30 frames per second, each of which could potentially contain political content in isolation from other frames. The framerate of the data collection application in this dissertation is 0.2 frames per second, meaning that a resolution 150 more granular is possible on most devices. If the exponential curve of segment duration distribution continues its curve upward at the axis below the minimum detectable duration in this sample, then there is a possibility that my analysis was unable to capture a staggeringly large amount of extremely fleeting encounters. Future research could investigate what occurs at the smallest of intervals, and to what extent they impact subjects' experience, with special attention to the psychological processes possible at such short timescales. At the smallest durations or the largest, the relationship between duration and user retention has yet to be concretely determined. My dissertation stops shy of suggesting how long is long enough. This is where I most hope that future researchers investigate. While each outcome likely requires a baseline duration – a point that motivates proper explication of timescales – there is a baseline level of information retention or emotional imprint that underlies all media engagement. Learning this baseline in the smartphone environment may require ingenuity in survey design, ecological momentary assessment, and reference back to pioneering studies of attention to media conducted in the laboratory.

In sum, modern communication technology affords incredibly fast exposure to political content, and to remain accurate, measurement must catch up. The news format for political content has survived in various forms through the advent of broadcast radio, broadcast television, cable television, the early web, and interactive *Web 2.0*. With each step, greater customization of not just content but of timing has altered what news looks like and what a dosage of news might contain. Smartphones, social media, and the rest of the emergent universe of hyper-fast media have ensured that *political content,* as a topic, has broken the mold of news formats that carried it for a century. The political content we are exposed to is best measured in seconds at a time, meaning that audiences' method for understanding the political world is butting up against the lower temporal bound of genuine cognitive engagement, and straddling into an entirely different type of processing with timescales all its own. An encounter with political content can still lead to a thirty-minute viewing session on television, or an hour-long radio program, but to move forward in the field, we must incorporate the flash of a meme, a comment scrolled-by in a feed, or the tenth glimpse of a campaign email sitting in our inbox. Each of these instances surrounds us in time, and the possible effect of each instance on our political experience — however slight — is likely determined as much by its temporal scale as by its message.

Supplementary Appendix

Appendix A. Subject 43 behavioral walkthrough

Appendix B. Daily Timeline Idiosyncrasy Plots

Appendix C. Limited sample demographic information

As a matter of policy, the company that was contracted for recruitment does not enable the collection of personally identifiable information and does not allow validation checks on the solicitation of email addresses from respondents. To accommodate this, subjects were recruited via their panel aggregation system, and mid-survey, eligible respondents were directed to our onboarding website. While every effort was made to encourage and incentivize respondents to provide a single accurate email address on both the recruitment survey and their installation of the data collection application, many provided inaccurate or alternative email addresses in either entry. One additional effort to combat this obstacle was the embedding of query strings into links provided to respondents. These embedded strings would auto-populate onboarding instructions in our onboarding website, providing a unique numerical link between application account names and survey metadata. However, subjects frequently ignored these prompts, entering alternative numeric values.

While none of these obstacles impeded data collection via screen recording or the collection of consent forms, they prevented the connection of subjects' screenomes to the data collected in the recruitment survey, apart from core screening criteria met by all subjects. Emails to subjects requesting alternative email addresses did not improve the number of matches between screenomes and survey responses. In a last effort, four undergraduate researchers were hired in early 2020 for the job of matching survey information to screenomes according to information gleaned from ethnographic examination of screen recordings. This effort uncovered some additional matches, but also suggested that some subjects responded to multiple surveys with different demographic information, an unanticipated problem that other researchers may now be aware of.

Appendix D. Smartphone app categorization schema

The smartphone application categorization schema can be accessed through this link. If you are reading a paper copy of this document, please contact muise.dan@gmail.com for access.

Appendix E. Cross-category pairwise segment duration intraindividual proportion comparisons

Category 1	Category 2	Duration	Mean 1	Mean 2	P value	Holm-adj. P	Significant?
social media	news	5	54.09	47.24	0.13	1	0
social media	news	10	18.4	23.01	0.77	1	0
social media	news	15	8.88	8.54	0.84	1	0
social media	news	20	7.97	7.03	0.71	1	0
social media	news	25	2.34	4.1	0.34	1	0
social media	news	30	1.68	1.81	0.81	1	0
social media	news	35	1.61	1.09	0.41	1	0
social media	news	40	1.4	1.08	0.53	1	0
social media	news	45	0.48	0.76	0.43	1	0
social media	news	50	0.4	0.7	0.33	1	0
social media	news	55	0.65	0.48	0.33	1	0
social media	news	60	0.42	0.39	0.99	1	0
social media	communication	5	54.09	56.34	0.56	1	0
social media	communication	10	18.4	21.3	0.9	1	0
social media	communication	15	8.88	6.86	0.02	1	0
social media	communication	20	7.97	5.43	0.01	1	0
social media	communication	25	2.34	2.39	0.3	1	0
social media	communication	30	1.68	0.83	0	0.37	0
social media	communication	35	1.61	0.71	0.03	1	0
social media	communication	40	1.4	0.91	0.12	1	0
social media	communication	45	0.48	0.37	0.16	1	0
social media	communication	50	0.4	0.14	0	0.36	0
social media	communication	55	0.65	1.76	0.87	1	0
social media	communication	60	0.42	0.67	0.72	1	0
social media	browser	5	54.09	49.51	0.07	1	0
social media	browser	10	18.4	19.88	0.94	1	0
social media	browser	15	8.88	8.97	0.75	1	0
social media	browser	20	7.97	5.9	0.39	1	0
social media	browser	25	2.34	3.08	0.51	1	0
social media	browser	30	1.68	3.74	0.06	1	0
social media	browser	35	1.61	2.26	0.2	1	0
social media	browser	40	1.4	1.45	0.44	1	0
social media	browser	45	0.48	0.95	0.56	1	0
social media	browser	50	0.4	0.61	0.38	1	0
social media	browser	55	0.65	0.66	0.24	1	0
social media	browser	60	0.42	0.31	0.91	1	0
social media	other	5	54.09	50.84	0.23	1	0
social media	other	10	18.4	16.62	0.47	1	0
social media	other	15	8.88	9.46	0.99	1	0

social media	other	20	7.97	6.45	0.45	1	0
social media	other	25	2.34	3.26	0.06	1	0
social media	other	30	1.68	2.83	0.22	1	0
social media	other	35	1.61	2.38	0.31	1	0
social media	other	40	1.4	2.26	0.5	1	0
social media	other	45	0.48	1.37	0.17	1	0
social media	other	50	0.4	0.43	0.48	1	0
social media	other	55	0.65	0.58	0.04	1	0
social media	other	60	0.42	0.61	0.21	1	0
Category 1	**Category 2**	**Duration**	**Mean 1**	**Mean 2**	**P value**	**Holm-adj. P**	**Significant?**
news	communication	5	47.24	56.34	0.08	1	0
news	communication	10	23.01	21.3	0.61	1	0
news	communication	15	8.54	6.86	0.13	1	0
news	communication	20	7.03	5.43	0.06	1	0
news	communication	25	4.1	2.39	0.09	1	0
news	communication	30	1.81	0.83	0.02	1	0
news	communication	35	1.09	0.71	0.51	1	0
news	communication	40	1.08	0.91	0.15	1	0
news	communication	45	0.76	0.37	0.08	1	0
news	communication	50	0.7	0.14	0	0.1	0
news	communication	55	0.48	1.76	0.4	1	0
news	communication	60	0.39	0.67	0.84	1	0
news	browser	5	47.24	49.51	0.41	1	0
news	browser	10	23.01	19.88	0.79	1	0
news	browser	15	8.54	8.97	0.75	1	0
news	browser	20	7.03	5.9	0.34	1	0
news	browser	25	4.1	3.08	0.47	1	0
news	browser	30	1.81	3.74	0.34	1	0
news	browser	35	1.09	2.26	0.12	1	0
news	browser	40	1.08	1.45	0.96	1	0
news	browser	45	0.76	0.95	0.81	1	0
news	browser	50	0.7	0.61	0.78	1	0
news	browser	55	0.48	0.66	0.92	1	0
news	browser	60	0.39	0.31	0.93	1	0
news	other	5	47.24	50.84	0.31	1	0
news	other	10	23.01	16.62	0.35	1	0
news	other	15	8.54	9.46	0.88	1	0
news	other	20	7.03	6.45	0.3	1	0
news	other	25	4.1	3.26	0.93	1	0
news	other	30	1.81	2.83	0.67	1	0
news	other	35	1.09	2.38	0.17	1	0
news	other	40	1.08	2.26	0.9	1	0
news	other	45	0.76	1.37	0.88	1	0

news	other	50	0.7	0.43	0.58	1	0
news	other	55	0.48	0.58	0.66	1	0
news	other	60	0.39	0.61	0.4	1	0
Category 1	**Category 2**	**Duration**	**Mean 1**	**Mean 2**	**P value**	**Holm-adj. P**	**Significant?**
communication	browser	5	56.34	49.51	0.02	1	0
communication	browser	10	21.3	19.88	0.77	1	0
communication	browser	15	6.86	8.97	0.08	1	0
communication	browser	20	5.43	5.9	0.09	1	0
communication	browser	25	2.39	3.08	0.1	1	0
communication	browser	30	0.83	3.74	0	0	1
communication	browser	35	0.71	2.26	0	0.03	1
communication	browser	40	0.91	1.45	0.02	1	0
communication	browser	45	0.37	0.95	0.04	1	0
communication	browser	50	0.14	0.61	0	0.02	1
communication	browser	55	1.76	0.66	0.28	1	0
communication	browser	60	0.67	0.31	0.82	1	0
communication	other	5	56.34	50.84	0.09	1	0
communication	other	10	21.3	16.62	0.52	1	0
communication	other	15	6.86	9.46	0.01	0.97	0
communication	other	20	5.43	6.45	0.04	1	0
communication	other	25	2.39	3.26	0	0.31	0
communication	other	30	0.83	2.83	0	0	1
communication	other	35	0.71	2.38	0	0.08	0
communication	other	40	0.91	2.26	0.02	1	0
communication	other	45	0.37	1.37	0	0.36	0
communication	other	50	0.14	0.43	0	0.02	1
communication	other	55	1.76	0.58	0.04	1	0
communication	other	60	0.67	0.61	0.09	1	0
Category 1	**Category 2**	**Duration**	**Mean 1**	**Mean 2**	**P value**	**Holm-adj. P**	**Significant?**
browser	other	5	49.51	50.84	0.42	1	0
browser	other	10	19.88	16.62	0.4	1	0
browser	other	15	8.97	9.46	0.68	1	0
browser	other	20	5.9	6.45	0.91	1	0
browser	other	25	3.08	3.26	0.27	1	0
browser	other	30	3.74	2.83	0.42	1	0
browser	other	35	2.26	2.38	0.67	1	0
browser	other	40	1.45	2.26	0.95	1	0
browser	other	45	0.95	1.37	0.42	1	0
browser	other	50	0.61	0.43	0.79	1	0
browser	other	55	0.66	0.58	0.36	1	0
browser	other	60	0.31	0.61	0.15	1	0

Appendix F. 443 intra-segment cross-category traversals

#	Segment duration, seconds	Category 1	Category 2	Category 3	Category 4	Category 5
1	15	browser	cash/survey			
2	10	browser	cash/survey			
3	15	browser	cash/survey			
4	40	browser	cash/survey			
5	10	browser	cash/survey			
6	105	browser	cash/survey			
7	15	browser	communication	search		
8	25	browser	communication			
9	20	browser	communication			
10	30	browser	communication			
11	100	browser	communication			
12	25	browser	communication			
13	15	browser	communication			
14	15	browser	communication			
15	300	browser	communication			
16	15	browser	communication			
17	330	browser	communication			
18	10	browser	communication			
19	35	browser	communication			
20	55	browser	communication			
21	15	browser	communication			
22	10	browser	communication			
23	10	browser	communication			
24	65	browser	communication			
25	40	browser	communication			
26	105	browser	communication			
27	15	browser	communication			
28	65	browser	communication			
29	25	browser	communication			
30	30	browser	communication			
31	90	browser	communication			
32	180	browser	entertainment			
33	15	browser	entertainment			
34	10	browser	entertainment			
35	15	browser	entertainment			
36	10	browser	entertainment			
37	45	browser	news	multimedia		
38	15	browser	news			

39	40	browser	news			
40	140	browser	other	communication		
41	45	browser	other	news		
42	35	browser	other			
43	15	browser	other			
44	35	browser	other			
45	655	browser	other			
46	135	browser	other			
47	35	browser	other			
48	140	browser	other			
49	15	browser	other			
50	20	browser	other			
51	30	browser	other			
52	40	browser	other			
53	25	browser	other			
54	10	browser	other			
55	25	browser	other			
56	90	browser	search			
57	15	browser	search			
58	15	browser	search			
59	30	browser	search			
60	35	browser	search			
61	115	browser	search			
62	180	browser	search			
63	10	browser	search			
64	55	browser	search			
65	35	browser	search			
66	45	browser	search			
67	120	browser	search			
68	40	browser	search			
69	30	browser	search			
70	25	browser	search			
71	65	browser	search			
72	25	browser	social media			
73	40	browser	social media			
74	15	browser	social media			
75	10	browser	social media			
76	85	browser	social media			
77	25	browser	social media			
#	**Segment duration, seconds**	**Category 1**	**Category 2**	**Category 3**	**Category 4**	**Category 5**
78	40	cash/survey	browser			
79	30	cash/survey	browser			

80	45	cash/survey	browser		
81	735	cash/survey	communication	entertainment	news
82	30	cash/survey	communication	social media	
83	35	cash/survey	communication		
84	45	cash/survey	communication		
85	45	cash/survey	communication		
86	40	cash/survey	communication		
87	35	cash/survey	communication		
88	30	cash/survey	communication		
89	40	cash/survey	communication		
90	40	cash/survey	communication		
91	35	cash/survey	communication		
92	45	cash/survey	communication		
93	45	cash/survey	communication		
94	40	cash/survey	communication		
95	40	cash/survey	communication		
96	25	cash/survey	communication		
97	40	cash/survey	entertainment		
98	10	cash/survey	entertainment		
99	20	cash/survey	entertainment		
100	85	cash/survey	news	tool	
101	10	cash/survey	news		
102	10	cash/survey	news		
103	75	cash/survey	other	game	tool
104	45	cash/survey	other	travel	
105	45	cash/survey	other		
106	45	cash/survey	other		
107	15	cash/survey	other		
108	40	cash/survey	other		
109	40	cash/survey	other		
110	10	cash/survey	other		
111	10	cash/survey	other		
112	45	cash/survey	other		
113	15	cash/survey	other		
114	35	cash/survey	other		
115	20	cash/survey	other		
116	45	cash/survey	other		
117	40	cash/survey	other		
118	45	cash/survey	other		
119	40	cash/survey	other		
120	45	cash/survey	other		
121	35	cash/survey	other		
122	45	cash/survey	other		

123	50	cash/survey	other
124	40	cash/survey	other
125	10	cash/survey	other
126	35	cash/survey	other
127	40	cash/survey	other
128	45	cash/survey	other
129	50	cash/survey	other
130	10	cash/survey	other
131	35	cash/survey	other
132	35	cash/survey	other
133	20	cash/survey	other
134	30	cash/survey	other
135	40	cash/survey	other
136	30	cash/survey	other
137	20	cash/survey	other
138	10	cash/survey	other
139	20	cash/survey	other
140	30	cash/survey	other
141	30	cash/survey	other
142	25	cash/survey	other
143	30	cash/survey	other
144	25	cash/survey	other
145	115	cash/survey	other
146	30	cash/survey	other
147	30	cash/survey	other
148	15	cash/survey	other
149	25	cash/survey	other
150	25	cash/survey	other
151	40	cash/survey	other
152	30	cash/survey	other
153	35	cash/survey	other
154	35	cash/survey	other
155	25	cash/survey	other
156	25	cash/survey	other
157	10	cash/survey	search
158	45	cash/survey	search
159	40	cash/survey	search
160	20	cash/survey	social media
161	35	cash/survey	social media
162	25	cash/survey	social media
163	40	cash/survey	social media
164	40	cash/survey	social media
165	30	cash/survey	social media

#	Segment duration, seconds	Category 1	Category 2	Category 3	Category 4	Category 5
166	60	cash/survey	social media			
167	70	communication	browser			
168	10	communication	browser			
169	20	communication	browser			
170	15	communication	browser			
171	45	communication	browser			
172	10	communication	browser			
173	15	communication	browser			
174	10	communication	browser			
175	10	communication	browser			
176	30	communication	browser			
177	25	communication	browser			
178	10	communication	browser			
179	20	communication	browser			
180	35	communication	browser			
181	20	communication	browser			
182	55	communication	browser			
183	20	communication	browser			
184	35	communication	browser			
185	30	communication	browser			
186	20	communication	cash/survey			
187	20	communication	cash/survey			
188	10	communication	cash/survey			
189	40	communication	cash/survey			
190	10	communication	entertainment			
191	10	communication	entertainment			
192	15	communication	entertainment			
193	20	communication	entertainment			
194	45	communication	entertainment			
195	40	communication	entertainment			
196	20	communication	news			
197	100	communication	news			
198	10	communication	news			
199	60	communication	news			
200	10	communication	news			
201	45	communication	news			
202	45	communication	news			
203	10	communication	news			
204	15	communication	news			
205	10	communication	news			
206	10	communication	news			

207	15	communication	other	entertainment		
208	775	communication	other	entertainment		
209	25	communication	other			
210	20	communication	other			
211	10	communication	other			
212	15	communication	other			
213	20	communication	other			
214	20	communication	other			
215	55	communication	other			
216	15	communication	other			
217	80	communication	other			
218	10	communication	search			
219	30	communication	search			
220	45	communication	search			
221	15	communication	search			
222	10	communication	search			
223	10	communication	social media			
224	65	communication	social media			
225	25	communication	social media			
226	95	communication	social media			
227	85	communication	social media			
228	10	communication	social media			
229	10	communication	social media			
230	10	communication	social media			
231	125	communication	social media			
#	**Segment duration, seconds**	**Category 1**	**Category 2**	**Category 3**	**Category 4**	**Category 5**
232	60	entertainment	browser	search		
233	110	entertainment	browser			
234	35	entertainment	browser			
235	550	entertainment	browser			
236	190	entertainment	cash/survey	finance		
237	45	entertainment	cash/survey			
238	40	entertainment	cash/survey			
239	30	entertainment	cash/survey			
240	70	entertainment	cash/survey			
241	30	entertainment	cash/survey			
242	70	entertainment	communication			
243	80	entertainment	communication			
244	95	entertainment	communication			
245	90	entertainment	communication			
246	55	entertainment	communication			
247	15	entertainment	communication			

248	15	entertainment	communication			
249	40	entertainment	news	cash/survey		
250	50	entertainment	news			
251	20	entertainment	news			
252	10	entertainment	news			
253	170	entertainment	news			
254	755	entertainment	other	news		
255	15	entertainment	other			
256	790	entertainment	other			
257	30	entertainment	other			
258	25	entertainment	other			
259	255	entertainment	other			
260	10	entertainment	other			
261	15	entertainment	social media			
262	15	entertainment	social media			
#	**Segment duration, seconds**	**Category 1**	**Category 2**	**Category 3**	**Category 4**	**Category 5**
263	45	news	browser	lifestyle		
264	15	news	browser			
265	10	news	browser			
266	310	news	cash/survey	health	entertainment	
267	10	news	cash/survey			
268	10	news	cash/survey			
269	35	news	cash/survey			
270	15	news	cash/survey			
271	10	news	communication			
272	45	news	communication			
273	55	news	communication			
274	80	news	communication			
275	15	news	communication			
276	100	news	communication			
277	65	news	communication			
278	45	news	communication			
279	25	news	communication			
280	40	news	communication			
281	25	news	communication			
282	85	news	communication			
283	25	news	communication			
284	15	news	communication			
285	25	news	entertainment	health		
286	80	news	entertainment			
287	15	news	entertainment			
288	240	news	entertainment			

289	20	news	entertainment			
290	35	news	other	health		
291	10	news	other			
292	10	news	other			
293	50	news	other			
294	15	news	other			
295	55	news	other			
296	10	news	other			
297	20	news	search			
298	80	news	search			
299	10	news	search			
300	275	news	social media	travel		
#	**Segment duration, seconds**	**Category 1**	**Category 2**	**Category 3**	**Category 4**	**Category 5**
301	90	other	browser			
302	10	other	browser			
303	10	other	browser			
304	25	other	browser			
305	300	other	cash/survey	communication		
306	30	other	cash/survey	health		
307	10	other	cash/survey			
308	20	other	cash/survey			
309	25	other	cash/survey			
310	40	other	cash/survey			
311	10	other	cash/survey			
312	15	other	cash/survey			
313	125	other	cash/survey			
314	15	other	cash/survey			
315	40	other	cash/survey			
316	105	other	communication	cash/survey	finance	social media
317	370	other	communication	news	entertainment	
318	470	other	communication	news		
319	215	other	communication	tool		
320	10	other	communication			
321	25	other	communication			
322	60	other	communication			
323	10	other	communication			
324	10	other	communication			
325	10	other	communication			
326	40	other	communication			
327	80	other	communication			
328	20	other	communication			
329	25	other	communication			

330	35	other	communication			
331	165	other	communication			
332	55	other	entertainment	tool		
333	95	other	entertainment			
334	65	other	entertainment			
335	15	other	entertainment			
336	10	other	entertainment			
337	1780	other	news	game	cash/survey	browser
338	10	other	news			
339	15	other	news			
340	10	other	news			
341	10	other	news			
342	25	other	news			
343	25	other	other	browser		
344	375	other	other	news		
345	25	other	other			
346	15	other	other			
347	20	other	other			
348	45	other	other			
349	60	other	other			
350	15	other	other			
351	60	other	other			
352	25	other	other			
353	70	other	other			
354	25	other	other			
355	360	other	other			
356	15	other	other			
357	10	other	other			
358	15	other	other			
359	10	other	search			
360	15	other	search			
361	130	other	social media			
362	15	other	social media			
363	50	other	social media			
364	55	other	social media			
365	10	other	social media			
366	25	other	social media			
#	**Segment duration, seconds**	**Category 1**	**Category 2**	**Category 3**	**Category 4**	**Category 5**
367	165	search	browser	communication		
368	20	search	browser			
369	65	search	browser			
370	15	search	browser			

371	20	search	browser			
372	340	search	browser			
373	100	search	browser			
374	240	search	browser			
375	30	search	browser			
376	25	search	browser			
377	35	search	browser			
378	30	search	browser			
379	20	search	browser			
380	15	search	browser			
381	90	search	browser			
382	15	search	browser			
383	10	search	browser			
384	30	search	browser			
385	15	search	browser			
386	85	search	browser			
387	20	search	browser			
388	90	search	browser			
389	30	search	browser			
390	55	search	browser			
391	60	search	browser			
392	35	search	browser			
393	60	search	communication	tool		
394	25	search	communication			
395	10	search	communication			
396	50	search	entertainment			
397	20	search	entertainment			
398	50	search	entertainment			
399	10	search	other			
#	**Segment duration, seconds**	**Category 1**	**Category 2**	**Category 3**	**Category 4**	**Category 5**
400	10	social media	browser			
401	65	social media	browser			
402	20	social media	browser			
403	20	social media	browser			
404	35	social media	browser			
405	10	social media	browser			
406	10	social media	browser			
407	70	social media	browser			
408	45	social media	cash/survey	tool		
409	35	social media	cash/survey			
410	15	social media	communication			
411	85	social media	communication			

412	20	social media	communication	
413	100	social media	communication	
414	20	social media	communication	
415	40	social media	communication	
416	20	social media	communication	
417	10	social media	communication	
418	15	social media	communication	
419	105	social media	communication	
420	25	social media	communication	
421	20	social media	communication	
422	10	social media	communication	
423	30	social media	communication	
424	15	social media	communication	
425	220	social media	entertainment	tool
426	45	social media	entertainment	
427	10	social media	entertainment	
428	25	social media	entertainment	
429	15	social media	entertainment	
430	30	social media	other	
431	35	social media	other	
432	15	social media	other	
433	50	social media	other	
434	20	social media	other	
435	15	social media	other	
436	10	social media	other	
437	55	social media	other	
438	130	social media	other	
439	25	social media	other	
440	35	social media	other	
441	105	social media	other	
442	20	social media	other	
443	10	social media	search	

Appendix G. Extended Analysis: valence, arousal, wordcount, and duration

In this appendix section, I plot out additional comparisons and descriptive analyses beyond what was shown or discussed in the final subsection of my analysis in Section 9. In Figure 18, I plot out multiple alternative comparisons using different parameters of valence, arousal and wordcount across segment durations. In Panels A, B, and C, extend Figure 17's Panel B. Specifically, I plot segment duration (in seconds, *x*-axis) against wordcount, arousal, and valence in time preceding the segment. In Panel A, this preceding duration is 30 seconds; in Panel B, this preceding duration is 60 seconds. In Panel C, I exactly replicate Panel B of Figure 17 by showing wordcount, valence, and arousal of the preceding five minutes. In the follow two rows of panels, the durations of 30, 60, and 300 seconds are the sole differentiator of the three columns' column content. In the second row, I extend Figure 17's Panel C by plotting differences of wordcount, valence, and arousal respectively between each segment and its preceding screenshots, against segment duration. Panel D compares intra-segment values against the preceding 30 seconds; Panel E compares intra-segment values against the preceding 60 seconds. In Panel F, I exactly replicate Panel C of Figure 17 by comparing against the preceding five minutes. In the third row, I add a wholly new component. Rather than examining what occurs *before* each segment, I show the valence, arousal, and wordcount in the screenshots *following* segments. Panel G, Panel H, and Panel I otherwise follow the exact same parameterization and axes of Panel A, Panel B, and Panel C of this figure, respectively. Across columns, Figure 18 suggests that variation in length of preceding (or succeeding) window of time abutting a segment does not substantively alter results. Similarity between Panel C and Panel I suggests that experience leading up to and following a segment is somewhat similar. Interestingly, Panel I shows that valence is significantly more negative following political content segments of any duration, though the substantive scale of this effect is small, and the reason is unclear.

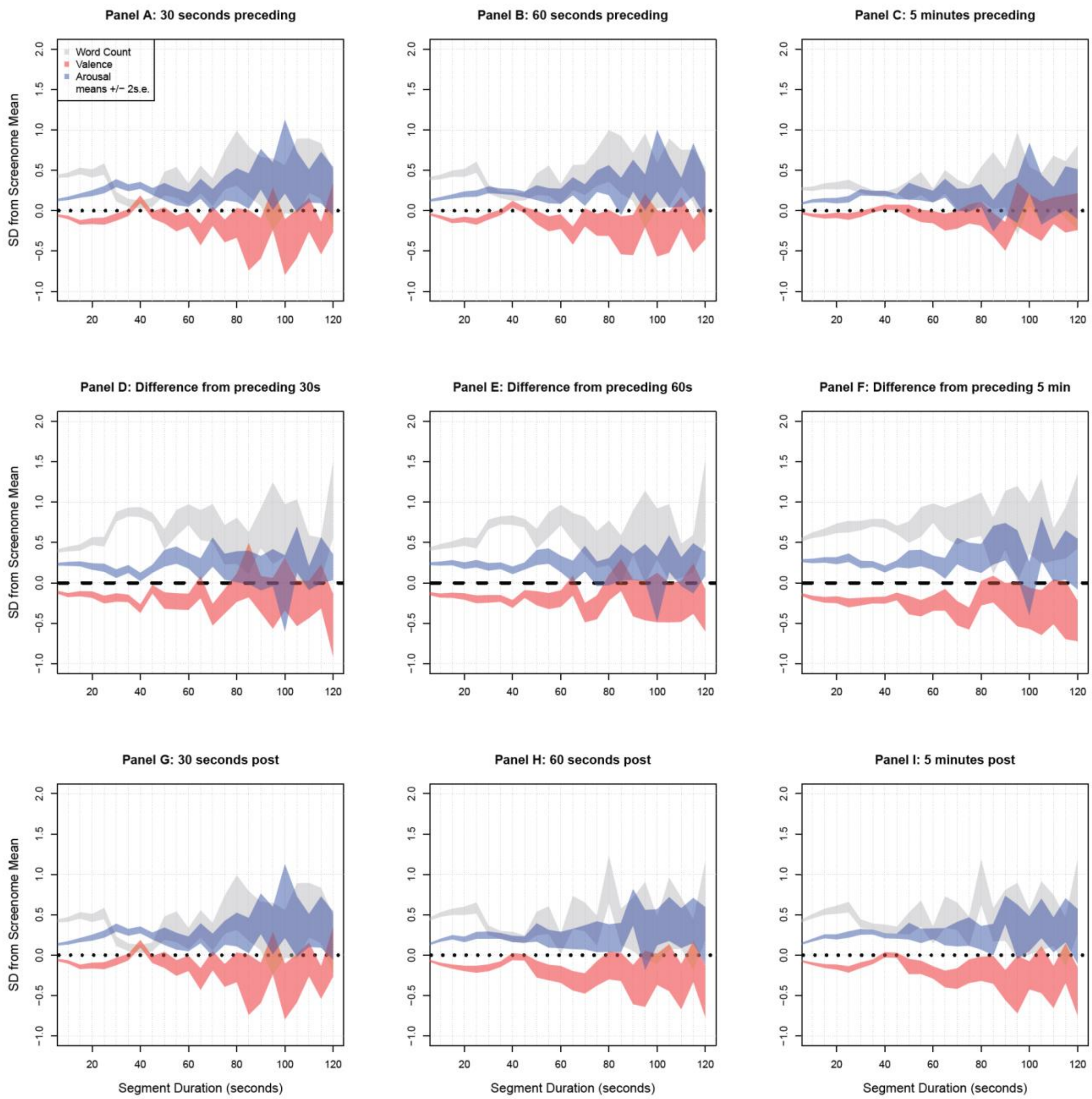
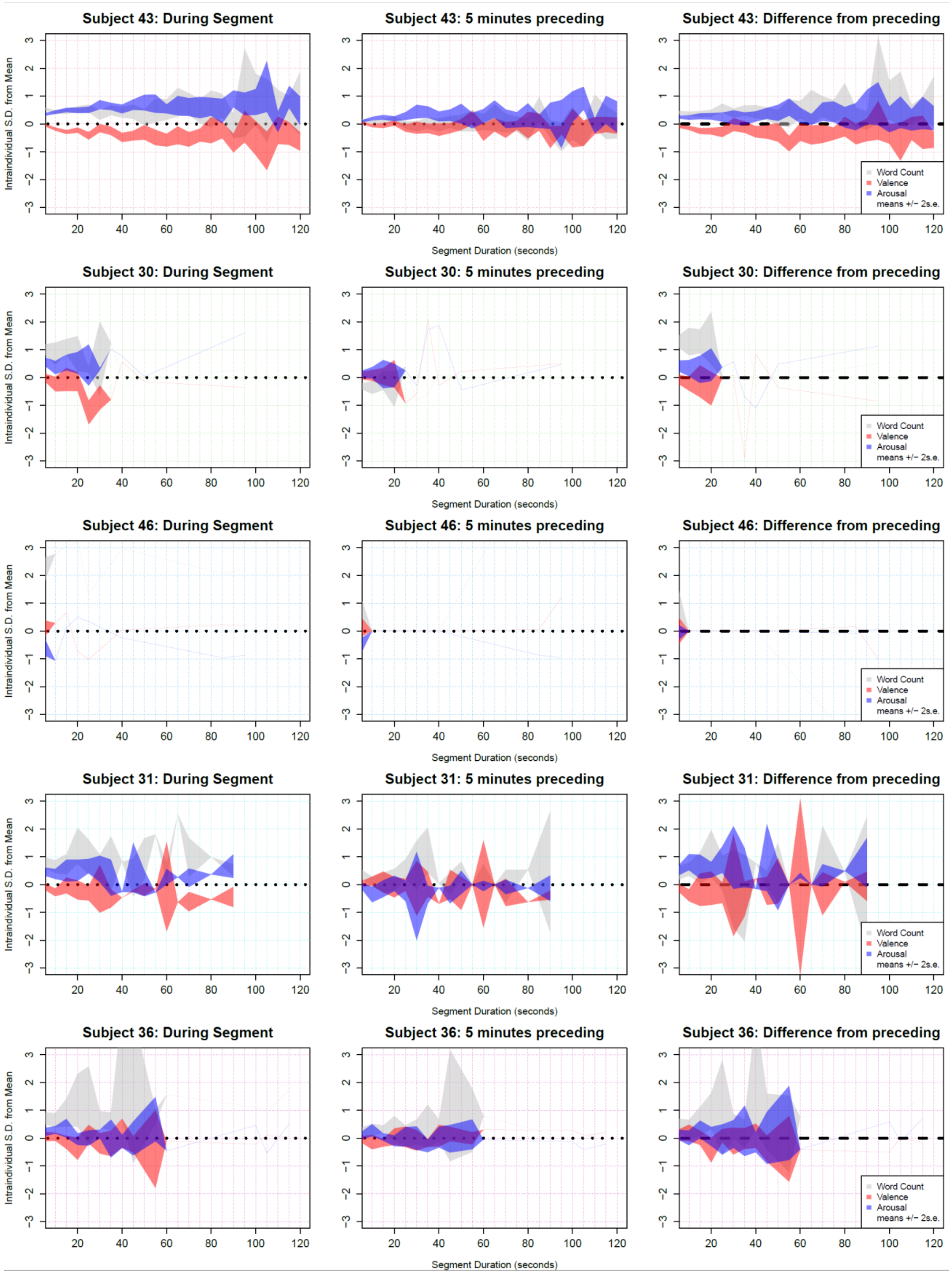


Figure 18. Segment duration versus wordcount, arousal, and valence during and preceding the segment. X-axes represent segment durations lasting between five and 120 seconds; y-axes represent intraindividual standard deviation from the intraindividual mean value. Bands represent pointwise mean values of all segments of duration x in the sample, +/- two pointwise standard errors. Overlap with the horizontal zero line indicates no significant difference from intra-individual mean value of each variable. Descriptive ordinary least square regression lines are overlaid in corresponding colors, with coefficient values and their significance level shown in-panel. Valence and arousal are based on NRC VAD lexicon (Mohammad, 2018). Segment durations describe whole segments, not

marginal increases. Panel columns represent preceding or succeeding periods lasting 30, 60, of 300 seconds respectively. Panel rows indicate three alternative analyses described by panel titles. See text for further clarification.

In Figure 19, I recreate Figure 17 intraindividually for the five familiar subjects referenced throughout the main analysis. Whereas bands in Figure 17 represented +/- two pointwise standard errors in interindividual mean estimates, bands in Figure 19 represent +/- two pointwise standard errors in mean estimates across aggregate pooling of segments in the sample. Out of these five subjects, only Subject 43 (see top panel row) experienced durations lasting up to 120 seconds. Subject 43's intraindividual experience mimics the interindividual experience visualized in Figure 17, indicating the strong of few subjects in describing behavior in longer-lasting segments, which may subjects do not encounter. In the remaining four rows of Figure 19, patterns found in the main analysis are more difficult to discern and less frequently statistically significant. However, the trend is similar, with valence being slightly more negative than average, arousal being slightly higher than average, and wordcount being slightly higher than average, relative each subject's average screenshot.



Subject 43: During Segment
Subject 43: 5 minutes preceding
Subject 43: Difference from preceding
Subject 30: During Segment
Subject 30: 5 minutes preceding
Subject 30: Difference from preceding
Subject 46: During Segment
Subject 46: 5 minutes preceding
Subject 46: Difference from preceding
Subject 31: During Segment
Subject 31: 5 minutes preceding
Subject 31: Difference from preceding
Subject 36: During Segment
Subject 36: 5 minutes preceding
Subject 36: Difference from preceding
Intraindividual S.D. from Mean
Segment Duration (seconds)
Word Count
Valence
Arousal
means +/− 2s.e.

Figure 19. For five subjects, referenced throughout the analysis: segment duration versus wordcount, arousal, and valence during and preceding the segment. X-axes represent segment durations lasting between five and 120 seconds; y-axes represent intraindividual standard deviation from the intraindividual mean value. Bands represent pointwise mean values of all segments of duration x in the sample, +/- two pointwise standard errors. Where two or fewer segments of X *duration occurred, bands are reduced to singular mean lines rather than expanding infinitely, for visual convenience. Overlap with the horizontal zero line indicates no significant difference from intra-individual mean value of each variable. Descriptive ordinary least square regression lines are overlaid in corresponding colors, with coefficient values and their significance level shown in-panel. Valence and arousal are based on NRC VAD lexicon (Mohammad, 2018).*

Appendix H. Politics classification details, extended

The keyword-based random forest method I use for classifying political screenshots is benchmarked against an initial bag of words approach. That bag of words approach employs a hybrid identification method that combines intuitive word selection with repeated manual auditing, similar to the approach used by Munger et al. (2017). Specifically, I start with a set of 152 word-stems linked to American political discourse in late 2019, which is the time period over which the data was collected. Initially, screenshots containing any single stem are treated as containing political content. The results were then compared to the ground truth set to detect initial accuracy metrics.

To improve precision in the bag of words approach, I manually examined screenshots in the training data which were incorrectly flagged as containing political content (false positives). I incrementally removed or down-weighted stems from the stem set that disproportionately drove the false positive rate. Some stems, such as *rally* and *debate*, are accurate in detecting political content only if in combination with at least one other stem, hence such stems were down-weighted to require the presence of another stem in the same screenshot. In the bag of words approach, applications that fall into the category of *game* were considered as never political. Many mobile games use war terminology on-screen during, e.g., sieging of castles or other fantasy scenarios far removed from political reality. Via manual inspection, no screenshot in my training data displayed a mobile game that included genuine political content. Finally, to reduce error introduced by the original app's foreground app detection method, no training screenshot with *Screenomics* as the foreground application was counted as containing political content.

The manual stem removal and down-weighting used in this baseline bag of words approach risks over-fitting to the training sample, despite the use of cross-validation. Thus, the bag of words approach is used only as a benchmark against which to measure a random forest based approach, and to that end, the potentially overfitted bag of words approach provides a conservative estimate of the added value of using the random forest approach. Manual omission or down-weighting of stems was not used in the random forest model used for classification in this dissertation. Rather, the model was fully nonparametric in its selection of stems to use as features, and in its assignment of the value of those stems in classification, thus mitigating overfit concerns.

Lastly, in Figure 20 I provide accuracy metrics capturing the performance of the random forest model using a range of probability thresholds. That is, the model used output probabilities $p \in [0,1]$ of a screenshot's likelihood of containing political content. I chose a threshold for positive classification by optimizing the F_1 metric, which properly weights performance in heavily unbalanced classification tasks. As shown, the optimal threshold value was the lowest non-zero probability value estimated by the model. As the stepwise probability output provided by the model resulted in the lowest step ranging $p \in (0,0.17]$, I used the mean value of this range as a threshold probability for the entire dataset. A worse-performing alternative random forest classification model based on text-embedding also required identifying an optimal positive classification threshold. Thus, the analogous F_1 optimization illustration for the embedding-based model is shown in Figure 21 immediately below Figure 20.

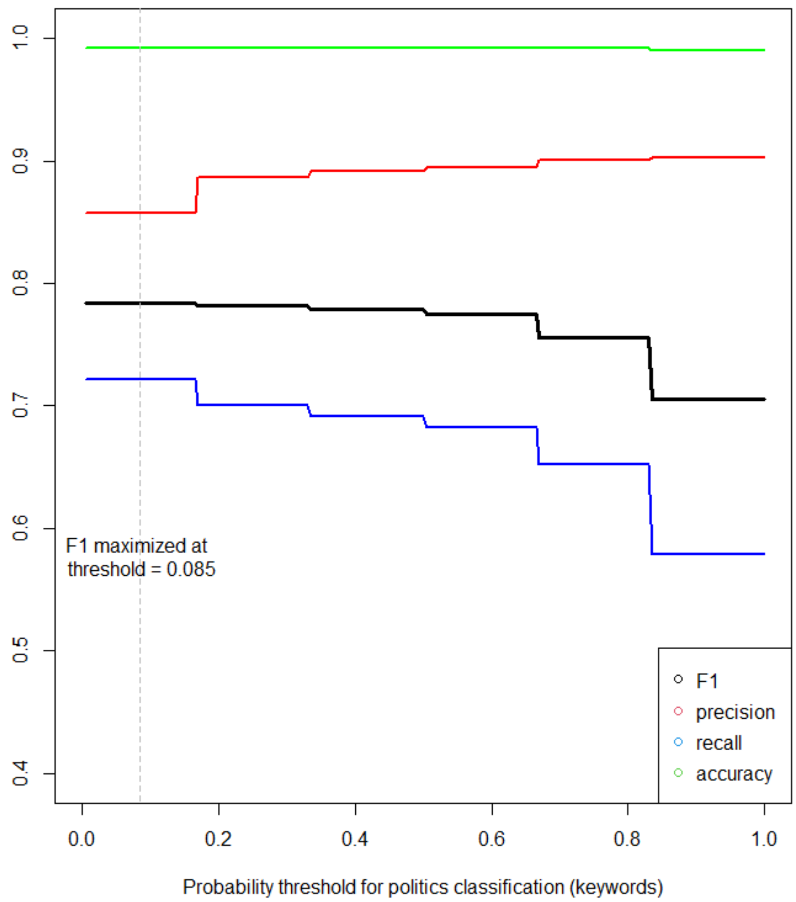


Figure 20. Four accuracy metrics describing the performance of my chosen keyword-based random forest model in the training data across variation in classification probability threshold. The chosen probability threshold, 0.85, was selected as it is the mean value of the F1-optimizing probability range.

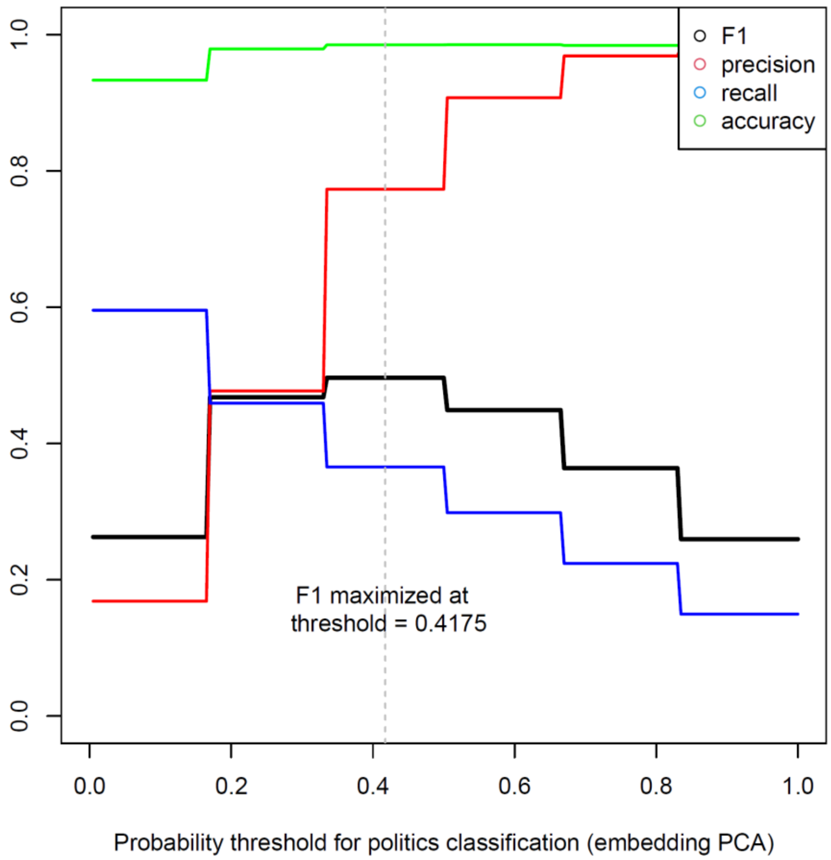


Figure 21. Four accuracy metrics describing the performance of the (not chosen text-embedding-based) random forest model in the training data across variation in classification probability threshold. The chosen probability threshold, 0.4175, was selected as it is the mean value of the F_1-optimizing probability range.

Appendix I. Robustness to survey-taking behavior

The sample recruited for this dissertation were found via a panel aggregator, and in the course of my analysis, I uncovered that subjects in my sample spent a large amount of time using application for earning cash incentives, either by completing surveys, performing tasks such as those on *Mechanical Turk*, and even gambling behavior. For some subjects, this category of applications which I refer to as *survey/cash* was the source of many segments of political content, and in a few cases, was a primary source. In the main text, I chose to analyze this behavior for what it is: natural smartphone usage for these subjects. However, in the interest of completeness, and to demonstrate robustness of core findings to the inclusion or exclusion of these applications, in this section I recreate all main text figures under the added condition that *survey/cash* applications are omitted. This was done by removing survey-taking applications from the screenome of each subject. What follows is each such figure, in order, captioned identically to the corresponding figure in the main text. As interpretations are largely the same as what was provided in the main analysis, I do not provide further figure discussion in this section.

	Total	Mean	Minimum	Median	Maximum
Subjects	115				
Screenshots	4,205,410	36,568	1,069	33,864	178,094
Screentime (hours)	5841	50.8	1.5	47	247.4
Subject Days	1,463	13	6	14	14
Daily Screentime (hours/day)		3.9	0.1	3.5	17.7

Table 7. Refers to Table 2. Description of the sample. Rows indicate sample features. Screenshots are the basic unit of data, though not always the unit of analysis in this dissertation. Screentime, shown in hours, is the approximate total screen activity of all subjects estimated through screenshot count. Subject-days is the number of days of smartphone usage included in the subsample. The sample is trimmed such that all subjects contribute a minimum of 6 consecutive days wherein the screen was never inactive for longer than 48 hours.

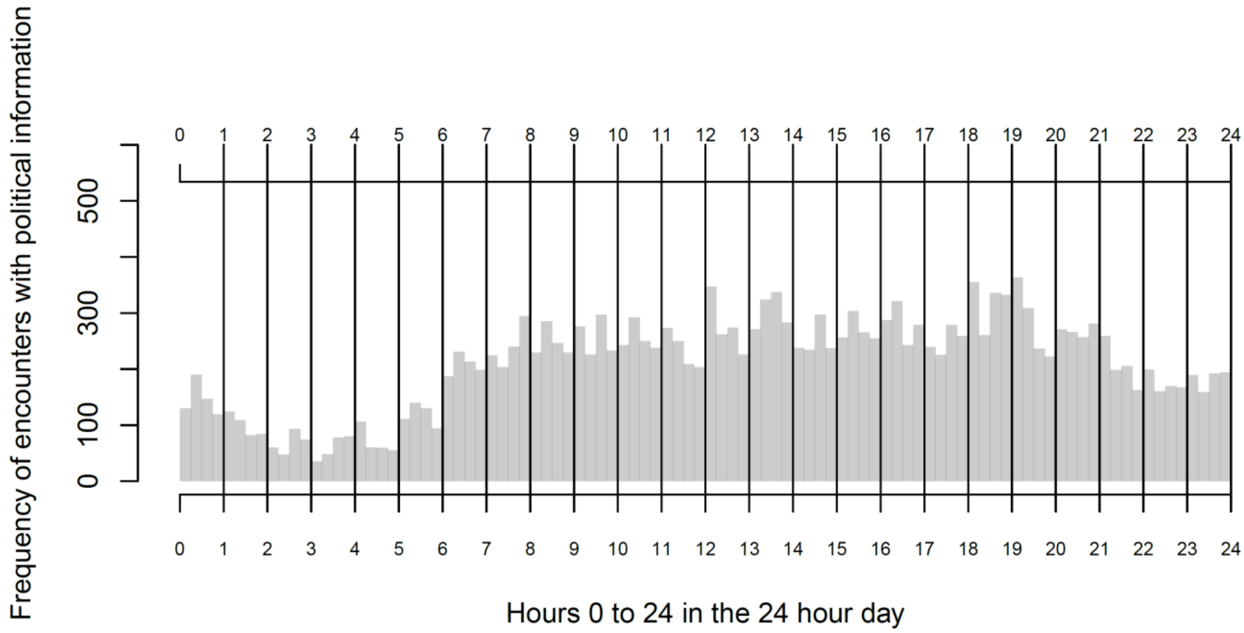


Figure 22. Refers to Figure 3. Political content encounters on the smartphone screen across 115 subjects (1463 subject-days) standardized to a single 24-hour clock, midnight to midnight. Encounters are bucketed into 15-minute blocks spanning the entire day. Note that 'encounter' refers to only the first screenshot in a series of continuous screenshots in which political content is identified, as explained in Figure 2.

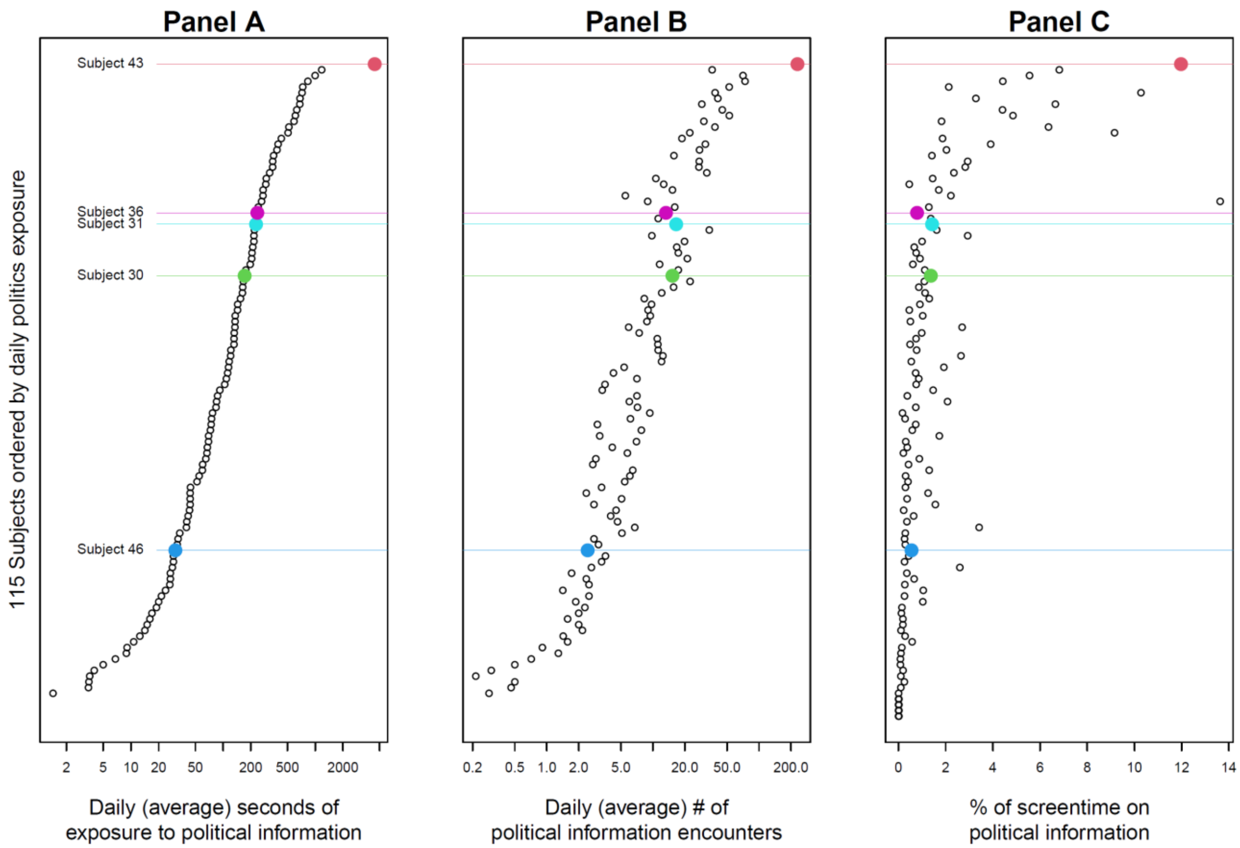


Figure 23. Refers to Figure 5. Inter-individual variation in exposure to political content in terms of daily seconds (Panel A), percentage of overall daily screentime (Panel B), and daily frequency of unique encounters (Panel C). Daily values are averages calculated across all unmanipulated subject-days, per subject. In Panel A, subjects are ordered according to their x-axis value; in Panels B and C, subjects maintain their ordering from Panel A. Note the log scales on Panels A and C, and the corresponding subjects shown from Figure 4.

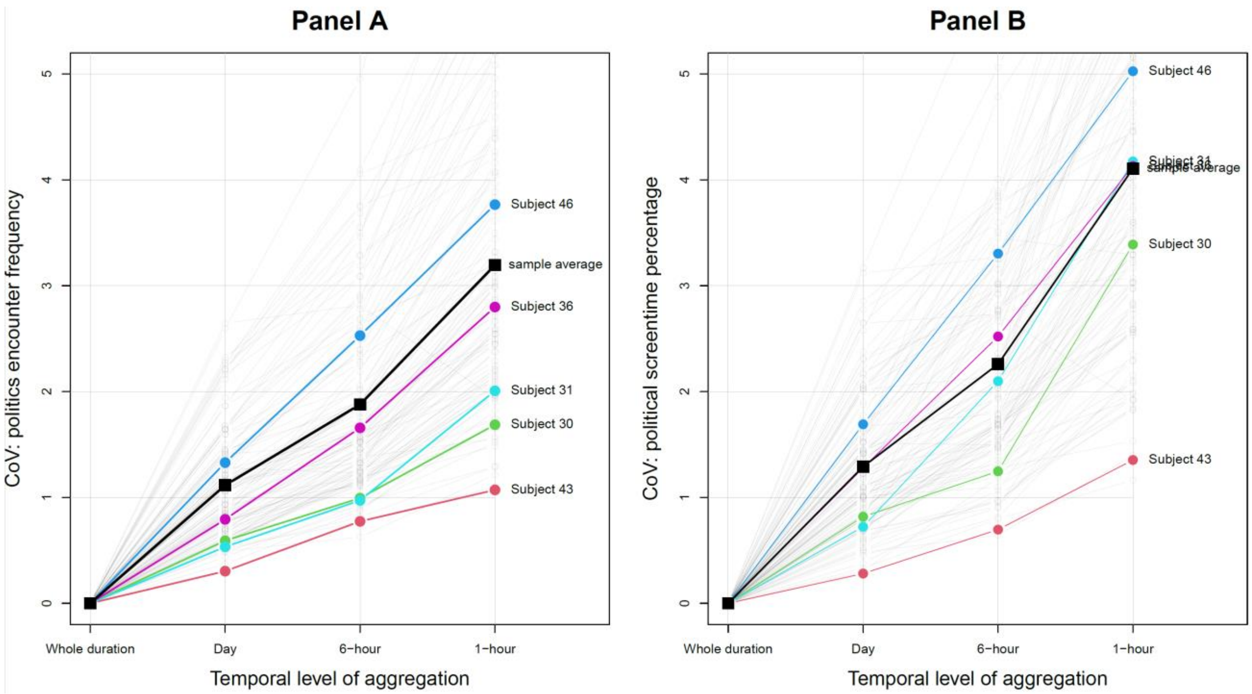


Figure 24. Refers to Figure 6. Coefficient of Variation (CoV) of daily frequency of political content encounters (Panel A) and the percentage of screentime during which political content was on-screen (Panel B). The y-axis shows CoV values ranging from 0 to 5, and the x-axis shows temporal levels of aggregation ranging from a subject's entire study duration to hour-long blocks at the most granular. All 115 subjects are shown in gray, with five subjects highlighted for reference, and sample-wide averages shown in bold black squares. On average across the sample, estimates of hourly politics encounter frequency misrepresent a randomly selected hour by a factor of 3.1, in expectation. On average across the sample, estimates of political screentime percentage misrepresent a randomly selected hour by a factor of 4, in expectation.

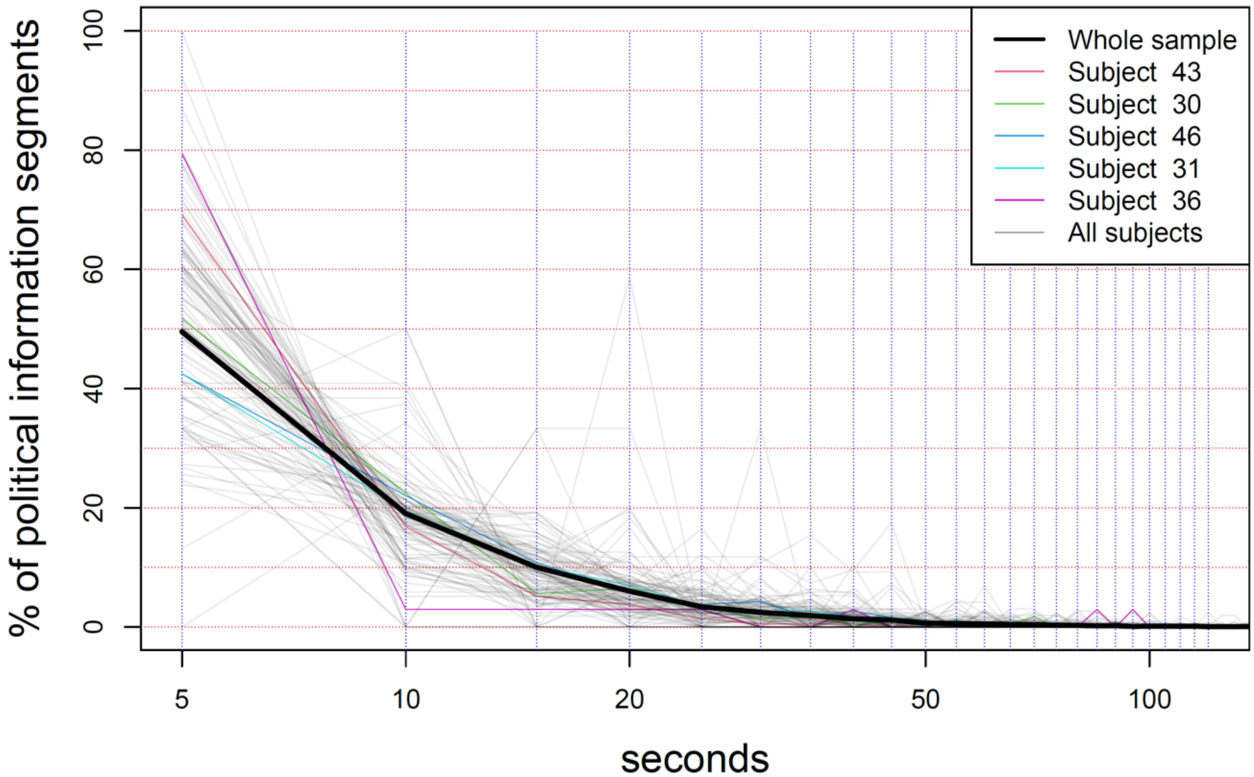


Figure 25. Refers to Figure 7. Distribution of political content segments across duration lengths. Segment length distribution for all segments in the data is shown in bold; each subject's individual distribution is underlaid in thin lines to shown inter-individual variation. Forty-four percent of all political content segments lasted just five seconds, the minimum detectable duration in the sample, and a duration at which recall and recognition rate is as low as 12% just minutes after exposure (Goldstein, McAfee, & Suri, 2011).

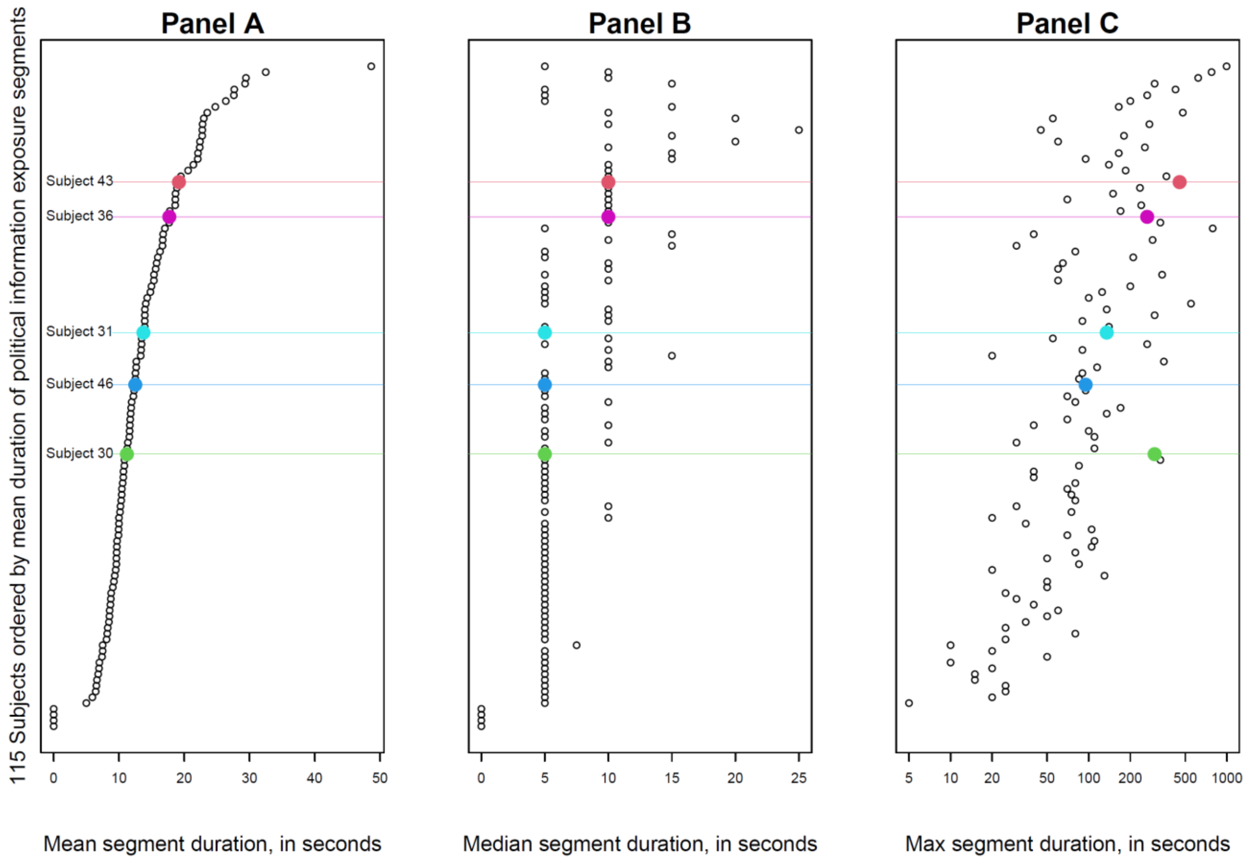


Figure 26. Refers to Figure 8. Mean duration (Panel A), median duration (Panel B), and maximum duration (Panel C) of political content segments for 115 subjects. Maximum duration is presented on a logarithmic scale. Subjects are ordered according to mean segment duration. Five subjects are highlighted for clarity. Most subjects' median segment duration was five seconds.

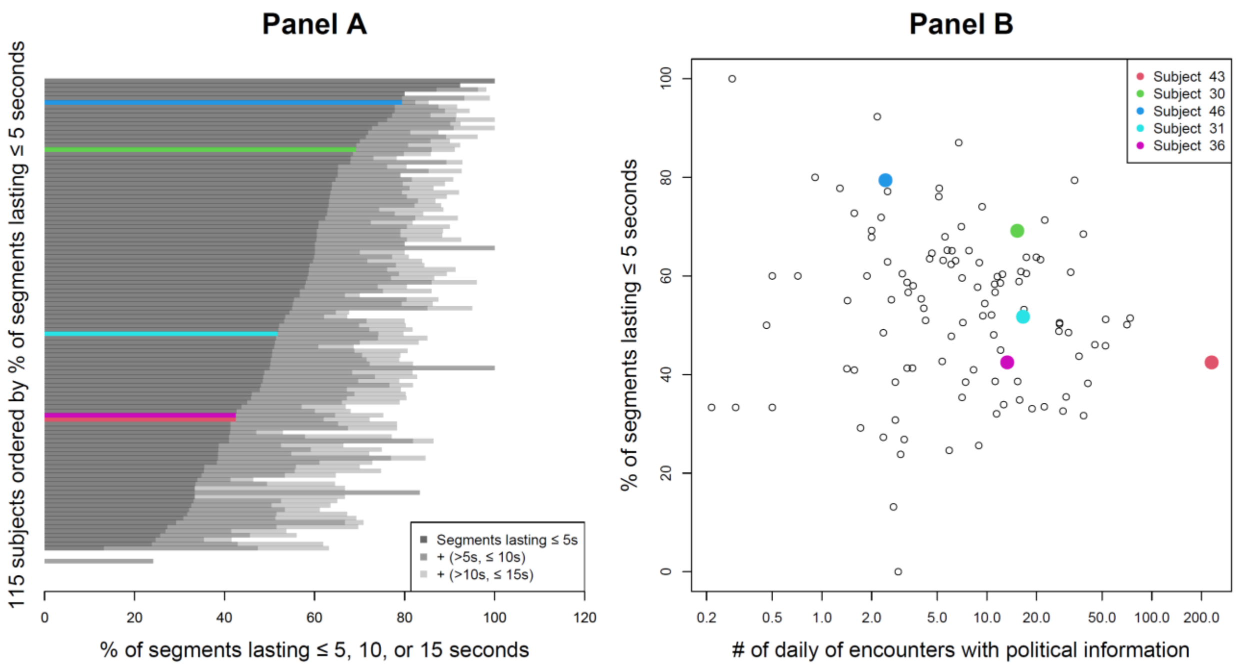


Figure 27. Refers to Figure 9. Subject-level segment durations. Panel A arranges 115 subjects according to the percent of their political content encounters that lead to 5-second segments. In progressively lighter shading, I show the additional percentage of encounters leading to 10-second-long and 15-second-long segments. In Panel B, I plot each subject according to the percent of their political content segments lasting five seconds (y-axis) and their number of daily encounters with political content. Note the logarithmic scale used in Panel B. Five subjects are highlighted for clarity.

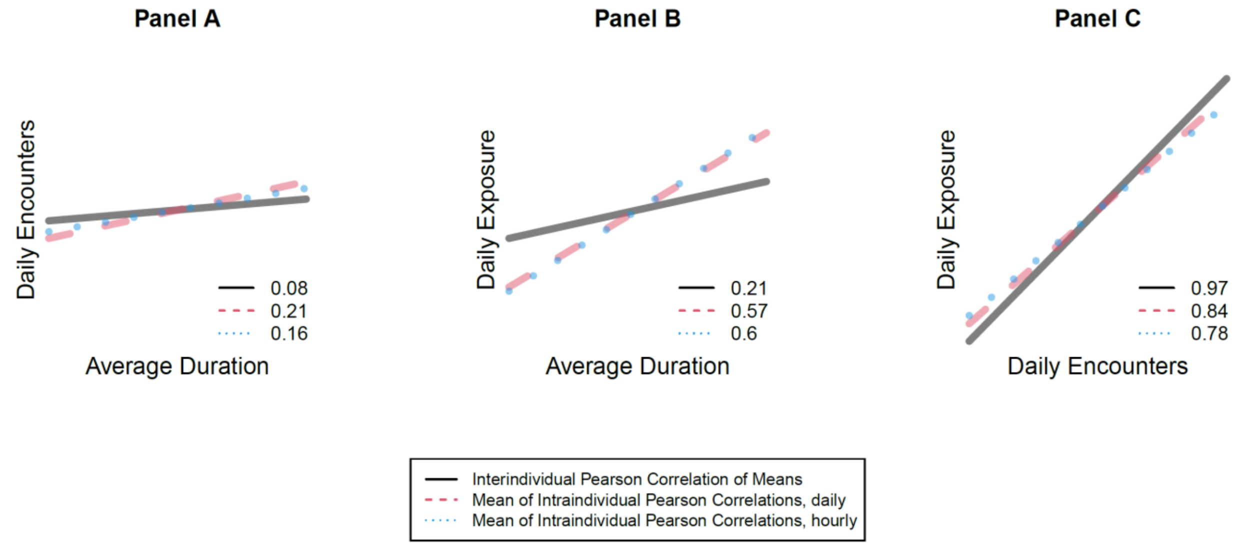


Figure 28. Refers to Figure 10. Interindividual versus intraindividual pairwise correlations of three variables summarizing political content exposure: average segment duration, number of encounters, and total exposure. Interindividual correlations are calculated from the subjects' average behavior in the data collection period. Intraindividual values are the average of intra-subject correlations between variables across days. Correlations are Pearson method, with standardization immediately prior to estimation. Correlations with opposing polarity represent non-ergodic pairwise relationships; all relationships display ergodicity.

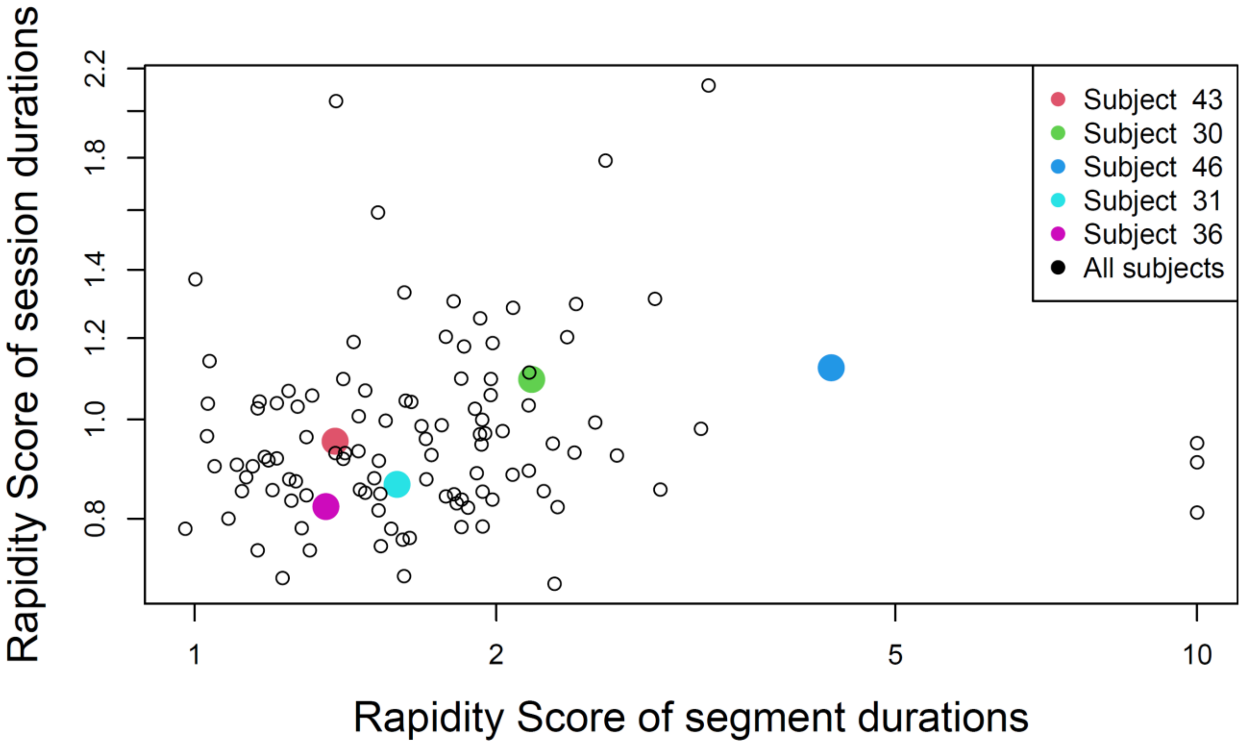


Figure 29. Refers to Figure 11. Subject-level scatter plot of rapidity scores of segment durations against rapidity scores of political content segment durations. A rapidity score is an MSE-minimizing degree describing the relationship between event and duration, where duration = (freq)$^{-degree}$. Higher rapidity scores implies greater tendency toward faster events versus longer ones. If a correlation exists between session rapidity and segment rapidity, then the lengths of subjects' segments might reflect underlying tendency toward longer durations of screen activity, or longer attention spans. However, there is no correlation. Subjects with no political content encounters are not included in this plot.

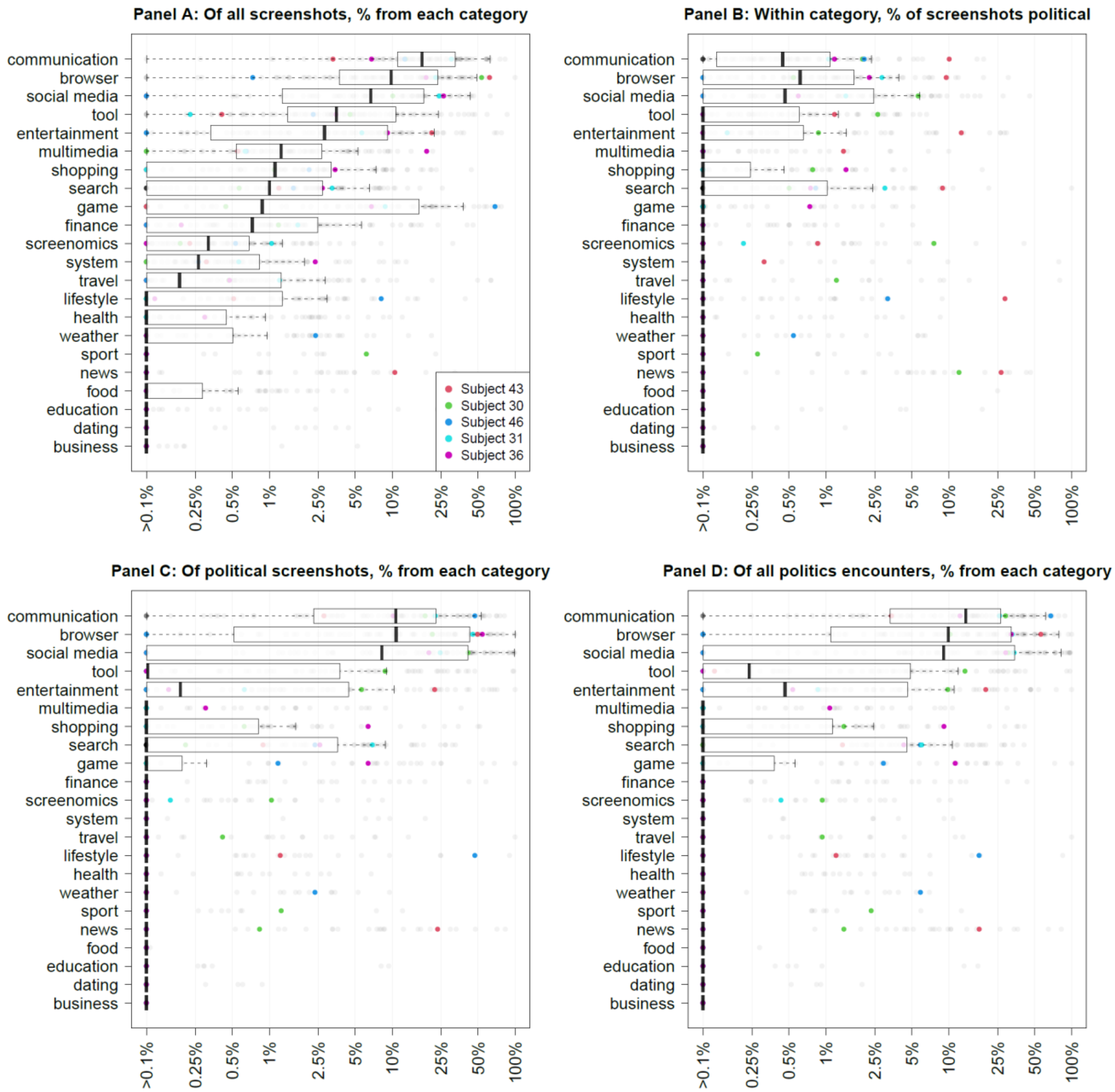


Figure 30. Refers to Figure 12. Decomposition of overall screen exposure and exposure to political content across 23 application categories for 109 subjects with foreground application data. Each subject is represented as a single point 23 times in each panel, with boxplots summarizing subjects' values within each application category. X-axes are explained in panel titles and shown in log scale; y-axis values are 23 categories in descending order according to medians in Panel A. Five subjects are highlighted in color for clarity and continuity with previous figures. Categories are based mainly on Google Play Store labels; see Appendix D.

Appendix J: Experimental Manipulation of Social Media Behavior

Here I describe a field experiment embedded in the data collection process of this dissertation. The experiment was underpowered due to challenges of implementation and attrition, and its foundational motivation — understanding political exposure in the smartphone environment — was more properly addressed by the observational analysis in the main text. The data I use in the main text are comprised of unmanipulated screenshots, that is, screenshots taken from the field experiment's control condition and the baseline period of the treatment group. The direct goal of the embedded field experiment was to measure the causal impact of a decrease in social media usage on exposure to political content, with the intent to better understand how partisan hostility may stem from social media usage. Similar analyses have recently been conducted by two highly powered field experiments focusing on the utility of Facebook, Instagram, and YouTube (Allcott, Gentzkow, Eichmeyer, & Braghieri, 2020; Aridor, 2022). Readers interested in causal effects in this domain are referred to those works. In the interest of completeness, this subsection details the design, implementation, and results of the embedded experiment. Results are inconclusive in that they are small, insignificant, trending in counterintuitive directions, and based on comparisons of cross-sectional samples insufficiently sized for drawing conclusions.

Using self-reports as outcome variables, Allcott et al. (2020) find a slight but significant decrease in polarization and demonstrated political knowledge following a four-week cessation of Facebook usage. The implication of this finding is that political content exposure on social media is (a) impactful on knowledge (b) not substitutionary with other platform sources and (c) hostility-inducing. A preregistration of the initially proposed field experiment (Muise, Pan, & Reeves, 2018) also aimed to measure pre/post changes in self-reported partisan hostility (see Appendix C for a description of low pre-survey data utility in the original data collection).

However, to the author's knowledge there is yet no *behavioral* evidence for a *causal* link between social media usage and political content exposure based on evidence from actual smartphone screens. As pertains specifically to this dissertation, this includes possible relationships between social media use and the frequency or duration of segments of political content. Based on that, a first hypothesis designed for behavioral data analysis was crafted prior to data collection which I adapt for my dissertation:

> H1: A decrease in social media usage decreases the frequency of encounters with political content.

Then, given the aims of my dissertation, here I add two additional behavioral hypotheses which were anticipated by the pre-registration as part of an exploratory analysis:

> H2: A decrease in social media usage decreases overall exposure to political content encounters.

> H3: A decrease in social media usage increases the average duration of political content segments.

Note that these hypotheses draw inspiration from theoretical aims outlined in the introduction of the main text. Note also that they are framed around a *decrease* in social media usage, rather than an increase. This matches the most practical and intuitive manner of intervening in subjects' social media usage, which is to reduce it either outright or in part.

Subjects were randomized into a treatment or control condition during the two-week study period. In the treatment condition, users were asked to avoid using *Facebook*, *Twitter*, *YouTube,* and *Reddit* for a period of three days. This three-day treatment period began in the middle of the two-week data collection period, typically on day 8, 9 or 10. As social media avoidance is a potentially demanding request for subjects, and one that is valued differently across subjects, treatment assignment followed a complex two-step design also used by Allcott et al. (2020). Our treatment assignment is designed according to the Becker-DeGroot-Marschak mechanism (BDM) (Becker, DeGroot and Marschak, 1964), which collects information on each subject's willingness to participate in the treatment condition. In this study, each individual was presented with an identical prompt, soliciting the minimum cash amount they would accept in exchange for complying with the treatment. One half of the participants were randomized into receiving a version of the prompt which offers $35 in exchange for treatment compliance, as long as their personal valuation was $35 or less. To implement the WTA collection and treatment assignment, all subjects received an email message in the middle of the data collection period soliciting them to take a short survey worth $5. All of these surveys contained identical information, and introduced the treatment parameters uniformly to all participants, soliciting the WTA. Participants in the control group received a survey for which the BDM's predetermined offer value is $0, and hence the control group participants could not declare a WTA low enough to be offered to join the treatment. Those in the treatment group who expressed a WTA of $35 or less were shown these exact instructions:

> We are offering you $35 to stop using Facebook, Reddit, YouTube, and Twitter from 12:01am [tomorrow's date] until 11:59pm [3 days later] (Pacific Time).

Dates and time zones were filled-in automatically based on the subject's local time. Those who received the offer were able to accept or decline. Those who accepted agreed to comply with the instructions in exchange for receiving the additional $35 of compensation at the end of the data collection period.

To check for subjects' continued participation, I conducted periodic inspection of all subjects' most recent screenshot upload. Treatment participants were checked daily for compliance by checks of their application log data, specifically regarding the most recent instance of any of the four banned applications being used, and how long that usage session lasted. A prompt email was sent to any treatment condition subject for whom any of the four prohibited applications were detected during the treatment period. Each such email informed the subject about that the usage was detected by our system, allowed subjects the opportunity to explain themselves, and suggested three of the following responses: if they would like to restart the treatment period, if they would like to exit the treatment condition, or if they believe that they actually complied. Further action was dependent on the participants' response.[22]

Notwithstanding the availability of pre-survey data, 53 out of 115 subjects met the core criteria to be part of the experiment by signaling a Willingness To Accept (WTA) of $35 or less. The treatment condition contained 25 subjects, and the control condition contained 28. Out of 25 subjects in the treatment condition, 10 strictly complied with the treatment instructions by not using any prohibited applications during the treatment period. However, 5 of these 10 did not use any of the prohibited applications in the pre-treatment period either, leaving 5 subjects who complied with the treatment in the strictly expected manner. An additional 8 subjects in the treatment group partially complied, in that they used the prohibited applications for only 5 minutes per day on average during the treatment period. All 8 control group subjects who did not

[22] I would periodically run a custom script through the command line interface of the cloud account used in data collection. This script would produce a comma-separated values file listing the most recent instance of YouTube, Facebook, Reddit, or Twitter usage by each subject. Subjects who did not respond or chose to drop-out are regarded as non-compliers and removed from the treatment analysis. Subjects who provided a reasonable explanation for their social media usage (e.g., a momentary negligent lapse that they quickly terminated) or a reasonable protestation suggesting that the foreground app log was in error (e.g., a notification appeared on screen) were considered part of the treatment condition, provided than no further lapses occurred.

use any prohibited applications in the treatment period also did not use the prohibited applications in the pre-treatment period. Four subjects in the control condition were exposed to no political content in the pre-treatment period. The subset of subjects meeting all four stricter criteria (WTA $\leq$ \$35, exposure to political content in pre-treatment, usage of any prohibited application pre-treatment, and compliance with treatment for those in the treatment group) contains 20 control subjects and 5 treatment subjects.[23]

In Figure 31, I answer each of the three hypotheses raised above and provide a compliance check, using both the full experimental sample and the subset with ideal compliance. All comparisons shown in Figure 31 are the difference in difference measures. Specifically, for each of four variables, I calculate the intraindividual difference between treatment period behavior and pretreatment period behavior, and then run a standard *t* test of the resulting values between all subjects in the control condition and all subjects in the treatment condition. The points shown in Figure 31 are the difference between the two group-level averages per variable per subset, and wings are 95% confidence intervals on the estimate. Solid bands indicate that the test was conducted on the entire set of 53 experimental subjects, and dashed lines indicate the use of the stricter compliant subset. The dotted vertical line in red indicates no significant difference between treatment and control groups; where wings cross this line, the null hypothesis of zero difference between groups cannot be rejected.

[23] In comparison, prior to launching the data collection, I conducted a power analysis to determine the cross-sectional sample size needed for measurement of an expected effect size in opinion change. This power analysis determined that a cross-sectional sample size on the order of 200 subjects per condition would be required.

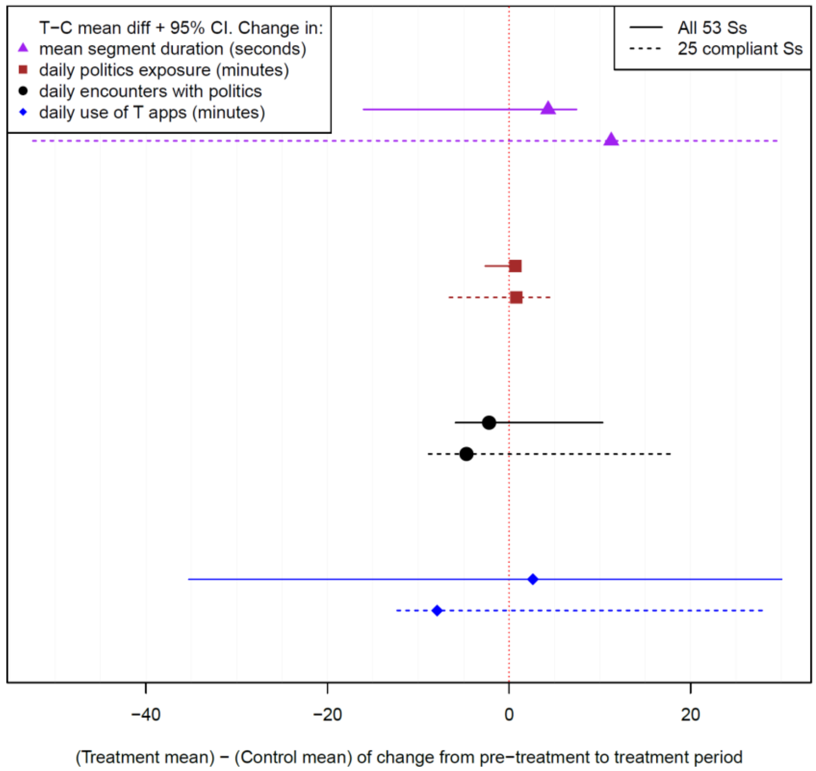


Figure 31. Compliance check and hypothesis testing results of embedded field experiment. In total, 28 out of 53 experiment-eligible subjects were randomly assigned to cease using Twitter, Facebook, Reddit, and YouTube (the "T applications") on their smartphones for three days after a week of unmanipulated smartphone usage. Variables are arranged from bottom to top to address a compliance check, H1, H2, and H3 (explained in-text). Variables are labeled in the top-left legend. Results from fully compliant subjects (those who used T applications prior to treatment period, were exposed to political content prior to the treatment period, and were compliant with the treatment) are shown in dashed lines (20C, 5T). All statistics are generated from standard difference-in-difference t*-tests.*

First, in blue diamonds, I provide a compliance check by examining the difference in change in daily usage in the four prohibited applications. For both the overall sample and the compliant subset, the difference between groups is insignificant, and in the full experimental sample, the difference is in a counterintuitive direction. It is possible that control condition subjects reduced their behavior in response to the solicitation of their WTA. This is not a possibility explored in applications of the WTA in similar research but may be a plausible explanation for the decrease in daily usage seen within some control group subjects.

Second, in black circles, I test Hypothesis 1. Hypothesis 1 was not supported, as the estimated changed in political content encounters was not significantly different between treatment and control groups. This implies that treatment subjects who reduced their social media usage still encountered political content as often as could be expected without any intervention. Had this result been found within a sample an order of magnitude larger, the interpretation would be that social media was a *substitute* to other sources of political content encounters on the smartphone, not a *complement*. Estimates trend in an expected direction for both the full experimental set and the subset of compliant subjects and are both very near to zero.

Third, in maroon squares, I test Hypothesis 2. Hypothesis 2 was not supported, as the estimated change in political content exposure was not significantly different between treatment and control groups. This implies that treatment subjects who reduced their social media usage still were exposed to political content for as many minutes per day could be expected without any intervention. Had this result been found within a sample an order of magnitude larger, the interpretation would be that social media was a *substitute* to other sources of political content exposure on the smartphone, not a *complement*. Estimates do trend in an unexpected direction for both the full experimental set and the subset of compliant subjects.

Fourth, in purple triangles squares, I test Hypothesis 3. Hypothesis 3 was not supported, as the estimated changed in average durations of political content segments was not significantly different between treatment and control groups. This implies that treatment subjects who reduced their social media usage still were exposed to political content segments lasting for as seconds on average as could be expected without any intervention. Had this result been found within a sample an order of magnitude larger, the interpretation would be that social media is a source of faster information segments than other sources of political content exposure on the smartphone, validated concerns about the likely lower quality of engagement political content often attributed to social media. Estimates do, however, trend in the expected direction for both the full experimental set and the subset of compliant subjects.

These results should not be interpreted as evidence of a lack of impact of social media use on political content exposure, either in quantity or temporal distribution. Rather, the sample used was insufficiently large (cross-sectionally) to properly identify the average treatment effect on the treated of decreasing social media usage. Other studies provide clearer guidance on the impact of a social media decrease on other domains of life. Allcott et al. (2021) found that total news exposure decreased as a result of decreased Facebook usage. They take this result to mean that news exposure is an externality of Facebook usage, such that subjects did not fully substitute their news exposure through other means. Allcott et al. (2021) also found individual-level political knowledge and policy issue polarization decreased following a decrease in Facebook usage, as indicated by self-reports, but not affective polarization or political engagement. This evidence points to a complex role of social media in causal political outcomes, meriting further exploration of individual encounters and segments.

Appendix K. Recruitment Survey

Stanford Screenomics Recruitment Survey

Start of Block: Stanford Screenomics

Stanford Screenomics
This is a recruitment survey for a research project, to measure how people use smartphones.

End of Block: Stanford Screenomics

Start of Block: .

If you're eligible, we will offer you $30 to install our app on your phone, for two weeks.

Our app is a scientific tool that records your device usage and sends that data to Stanford. We do not sell your data; we use this data for scientific research only. Would you consider joining this scientific project for $30?

○ I'm interested, tell me more

○ No, I'm not interested in this

Page Break

Display This Question:

If If you're eligible, we will offer you $30 to install our app on your phone, for two weeks. Our ap... = No, I'm not interested in this

Please describe why you are not interested:

____

Page Break

End of Block: .

Start of Block: Block 2

Great! First, we will ask a few questions to see if you are eligible for the study.

Page Break

Do you currently use only one smartphone daily, or multiple smartphones? (e.g., one for work and one for personal use)

- ○ I use one smartphone
- ○ I use multiple smartphones
- ○ I use zero smartphones on an average day

Skip To: End of Block If Do you currently use only one smartphone daily, or multiple smartphones? (e.g., one for work and... != I use one smartphone

Display This Question:

If Do you currently use only one smartphone daily, or multiple smartphones? (e.g., one for work and... = I use one smartphone

What is the brand of the smartphone you currently use?

- ◯ Apple (iPhone)
- ◯ Google (Pixel)
- ◯ HTC
- ◯ Huawei
- ◯ Lenovo
- ◯ LG
- ◯ Microsoft (Windows Phone)
- ◯ Motorola (Droid)
- ◯ Nokia (Blackberry)
- ◯ Samsung (Galaxy)
- ◯ Other ____

Skip To: End of Block If What is the brand of the smartphone you currently use? = Apple (iPhone)

Skip To: End of Block If What is the brand of the smartphone you currently use? = Microsoft (Windows Phone)

Display This Question:

If Do you currently use only one smartphone daily, or multiple smartphones? (e.g., one for work and... = I use multiple smartphones

What are the brands of the smartphones you currently use? Check all that apply.

- ▢ Apple (iPhone)
- ▢ Google (Pixel)
- ▢ HTC
- ▢ Huawei
- ▢ Lenovo
- ▢ LG
- ▢ Microsoft (Windows Phone)
- ▢ Motorola (Droid)
- ▢ Nokia (Blackberry)
- ▢ Samsung (Galaxy)
- ▢ Other ____

Page Break

Display This Question:

If Do you currently use only one smartphone daily, or multiple smartphones? (e.g., one for work and... = I use one smartphone

Are you the *only* adult who uses your smartphone?

◯ Yes, I am the only adult user of my smartphone

◯ No, my smartphone is also used by someone else/others

Skip To: End of Block If Are you the only adult who uses your smartphone? = No, my smartphone is also used by someone else/others

Page Break

Do any children use your smartphone? If so, when an how often?

◯ No children use my smartphone

◯ Yes, children use my smartphone. Here's often, what time of day, and for what purpose:
____

Page Break

Do you currently use only one computer/laptop daily, or do you use multiple computers? (e.g., one for work and one for personal use)

◯ I use one computer

◯ I use multiple computers

◯ I do not use a computer, just a smartphone

Page Break

Please describe up to three other electronic devices you own and use most often, such as tablets, smart watches, and televisions. Please include what you use the device for (e.g., web surfing, games) and how many hours per week you use these devices.

○ Additional Device 1 ____

○ Additional Device 2 ____

○ Additional Device 3 ____

Page Break

What is your year of birth?

▼ 2000 ... 1917

Display This Question:

If What is your year of birth? = 2000

Please confirm that you are over 18:

○ I am over 18 years old

○ I am not yet 18 years old

Skip To: End of Block If Please confirm that you are over 18: = I am not yet 18 years old

Page Break

Which of the following best describes you? (Choose one answer)

- Hispanic or Latino/Latina
- Not Hispanic or Latino/Latina

Page Break

Display This Question:

If Which of the following best describes you? (Choose one answer) = Not Hispanic or Latino/Latina

Which of the following best describes you? (Choose all that apply)

- American Indian or Alaskan Native
- Asian
- Black or African American
- Native Hawaiian or Pacific Islander
- White
- Other (please describe) ____
- I prefer not to answer

Page Break

What is your gender identity?

- ○ Male
- ○ Female
- ○ Other
- ○ I prefer not to answer

Page Break

What was your total household income from all sources before taxes last year? By "household", we mean that you should report the combined income of everyone in your home.

- ○ $14,999 or less
- ○ $15,000 - $24,999
- ○ $25,000 - $34,999
- ○ $35,000 - $49,999
- ○ $50,000 - $74,999
- ○ $75,000 - $99,999
- ○ $100,000 - $149,999
- ○ $150,000 - $199,999
- ○ $200,00 or more
- ○ Don't know
- ○ I prefer not to answer

Page Break

How many years of formal schooling do you have? E.g., years spent in 1st grade through 12th grade, plus years in college, plus years in graduate school.

▼ less than 12 ... 19+

Page Break

Now we would like to know something about your party preference and how you vote. Do you consider yourself a supporter of a particular political party? Which party?

- ◯ Republican
- ◯ Democrat
- ◯ Independent
- ◯ No preference
- ◯ Other ____

Page Break

Display This Question:

If Now we would like to know something about your party preference and how you vote. Do you consider... = Republican

Would you call yourself a strong Republican or a not very strong Republican?

- ◯ Strong
- ◯ Not very strong

Page Break

Display This Question:

If Now we would like to know something about your party preference and how you vote. Do you consider... = Democrat

Would you call yourself a strong Democrat or a not very strong Democrat?

- ◯ Strong
- ◯ Not very strong

Page Break

Display This Question:

If Now we would like to know something about your party preference and how you vote. Do you consider... = Independent

Or Now we would like to know something about your party preference and how you vote. Do you consider... = No preference

Or Now we would like to know something about your party preference and how you vote. Do you consider... = Other

Do you think of yourself as closer to the Republican Party or closer to the Democratic Party?

- ◯ Republican
- ◯ Democrat

End of Block: Block 2

Start of Block: Block 3

Now we will ask a few questions about you. The following questions do not affect your eligibility.

Page Break

What is your current marital status?

- ○ Married, or living as married
- ○ Divorced or separated
- ○ Widowed
- ○ Single/never married

Does anyone in your household receive food stamps or SNAP?

- ○ Yes
- ○ No
- ○ Don't know

Does anyone in your household receive Unemployment, Social Security, of Disability Benefits?

- ○ Yes
- ○ No
- ○ Don't know

Page Break

Does your family own the home in which you live?

- Yes
- No

What is your zip code?

____

How often do you speak English at home with your family?

- Never
- Sometimes
- About half of the time
- Most of the time
- Always

Page Break

How often do you use social media on an average day?

- not at all
- less than thirty minutes per day
- about one hour per day
- between 1 and 3 hours per day
- between 3 and 6 hours per day
- greater than 6 hours per day

Page Break

Some people don't pay much attention to politics. How about you, would you say that you have been/were very much interested, somewhat interested, or not much interested in following politics in the past year?

- Very much interested
- Somewhat interested
- Not much interested

Some people don't pay much attention to political *campaigns*. How about you? Would you say that you have been very much interested, somewhat interested or not much interested in the political *campaigns* so far this year?

- very much interested
- somewhat interested
- not very interested

Page Break

How often do you use these social media platforms?

	never	rarely	weekly	daily
Facebook	○	○	○	○
Instagram	○	○	○	○
Snapchat	○	○	○	○
YouTube	○	○	○	○
WeChat	○	○	○	○
Twitter	○	○	○	○
Reddit	○	○	○	○
Pinterest	○	○	○	○
tumblr	○	○	○	○
VKontakte	○	○	○	○
MySpace	○	○	○	○
LinkedIn	○	○	○	○
TikTok	○	○	○	○
Other:	○	○	○	○

Why do you use social media? Check all that apply:

- ☐ Keep up with politics
- ☐ Keep up with celebrities & entertainment
- ☐ Keep up with sports
- ☐ Keep up with local news & events
- ☐ See what friends & loved ones are up to
- ☐ Socialize with friends & loved ones
- ☐ For work and business
- ☐ Express my opinions
- ☐ Post updates about myself & my life, for my family
- ☐ Post updates about myself & my life, for my friends
- ☐ Post updates about myself & my life, to create my brand
- ☐ When I have nothing else to do / when I am bored
- ☐ Out of habit / unconsciously / I use social media without realizing
- ☐ Debate / discuss politics with others
- ☐ Look at entertaining videos / memes / pictures
- ☐ Play games

- Remember important events
- Keep track of important moments
- School / studying
- Get recommendations
- Meet new people / dating
- Join and participate in common-interest groups
- Look at / edit my own profile
- I don’t use social media
- None of these purposes/ other

Page Break

How do you normally read/watch/listen to political news? Check all that apply:

- ☐ TV
- ☐ Newspapers (on paper)
- ☐ Radio
- ☐ Laptop / tablet
- ☐ Smartphone

Display This Question:

If How do you normally read/watch/listen to political news? Check all that apply: = TV

Or How do you normally read/watch/listen to political news? Check all that apply: = Smartphone

Or How do you normally read/watch/listen to political news? Check all that apply: = Laptop / tablet

Which TV channel(s) do you normally watch?

____

Display This Question:

If How do you normally read/watch/listen to political news? Check all that apply: = TV

Or How do you normally read/watch/listen to political news? Check all that apply: = Laptop / tablet

Or How do you normally read/watch/listen to political news? Check all that apply: = Smartphone

Which TV program(s)/ TV host(s) do you normally watch?

____

Display This Question:

If How do you normally read/watch/listen to political news? Check all that apply: = Newspapers (on paper)

Or How do you normally read/watch/listen to political news? Check all that apply: = Laptop / tablet

Or How do you normally read/watch/listen to political news? Check all that apply: = Smartphone

Which newspaper(s) do you normally read?

____

Display This Question:

If How do you normally read/watch/listen to political news? Check all that apply: = Radio

Which radio station(s) do you normally listen to?

____

Display This Question:

If How do you normally read/watch/listen to political news? Check all that apply: = Radio

Which radio program(s) / radio host(s) do you normally listen to?

____

Page Break

Please look at the graphic below.

We'd like to know your feelings toward some political groups in the news these days. We'll show you the name of a group and we'd like you to rate that group using something we call the feeling thermometer. Ratings between 50 degrees-100 degrees mean that you feel favorable and warm toward the group; ratings between 0 and 50 degrees mean that you don't feel particularly favorable toward the group and that you don't care too much for that group. You would rate the group at the 50 degree mark if you don't feel particularly warm or cold toward the group. If there is a group that you don't recognize, you don't need to rate that group. Just click 'I don't recognize this group' and move on to the next one.

	unfavorable/cold favorable/warm I don't recognize this group. 0 10 20 30 40 50 60 70 80 90 100
Democratic Party	
Republican Party	
Conservatives	
Liberals	

Page Break

Display This Question:

If Now we would like to know something about your party preference and how you vote. Do you consider... = Democrat

Or Do you think of yourself as closer to the Republican Party or closer to the Democratic Party? = Democrat

How would you feel if you had a son or daughter who married a Republican supporter? Not at all upset, somewhat upset, very upset?

- ○ Not at all upset
- ○ Somewhat upset
- ○ Very upset

Display This Question:

If Now we would like to know something about your party preference and how you vote. Do you consider... = Republican

Or Do you think of yourself as closer to the Republican Party or closer to the Democratic Party? = Republican

How would you feel if you had a son or daughter who married a Democrat supporter? Not at all upset, somewhat upset, very upset?

- Not at all upset
- Somewhat upset
- Very upset

Page Break

Please rate your agreement with the following statements:

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
Sometimes I feel all alone in the world.	○	○	○	○	○
I sometimes feel uncertain about who I really am	○	○	○	○	○
Most people don't seem to accept me when I am just being myself.	○	○	○	○	○

Page Break

Display This Question:

If If To continue, open this Stanford URL: https://comm-research.stanford.edu. Then enter the four-digit passcode that is generated on that page here: Text Response Is Displayed

Congratulations, you are eligible to participate in the Stanford project! Click the red ">>" button below to proceed.

Skip To: End of Survey If Congratulations, you are eligible to participate in the Stanford project! Click the red " " butto... Is Displayed

To continue, open this Stanford URL: ${e://Field/rando}" target="_blank">https://comm-research.stanford.edu. Then enter the four-digit passcode that is generated on that page here:

____

Page Break

Display This Question:

If What is the brand of the smartphone you currently use? != Apple (iPhone)

And What is the brand of the smartphone you currently use? != Nokia (Blackberry)

And What is the brand of the smartphone you currently use? != Microsoft (Windows Phone)

And Do you currently use only one smartphone daily, or multiple smartphones? (e.g., one for work and... != I use multiple smartphones

And Are you the only adult who uses your smartphone? != No, my smartphone is also used by someone else/others

And Do you currently use only one computer/laptop daily, or do you use multiple computers? (e.g., one... != I use multiple computers

And Please confirm that you are over 18: != I am not yet 18 years old

Lastly, please enter your email address.
This helps us send reimbursement later (not for spam).

____

Display This Question:

If What is the brand of the smartphone you currently use? != Apple (iPhone)

And What is the brand of the smartphone you currently use? != Nokia (Blackberry)

And What is the brand of the smartphone you currently use? != Microsoft (Windows Phone)

And Do you currently use only one smartphone daily, or multiple smartphones? (e.g., one for work and... != I use multiple smartphones

And Are you the only adult who uses your smartphone? != No, my smartphone is also used by someone else/others

And Do you currently use only one computer/laptop daily, or do you use multiple computers? (e.g., one... != I use multiple computers

And Please confirm that you are over 18: != I am not yet 18 years old

You will now be redirected to a final page. You may close it and continue on the tab you just opened (${e://Field/rando}">https://comm-research.stanford.edu*)*

End of Block: Block 3

Appendix L. Consent Form

All research at Stanford University requires that we inform you about the study, tell you about how to contact people at Stanford University if you have any questions, and let you know that your participation is voluntary and that you may withdraw at any time. This is a form that has all of this information.

Please print or save a copy of this page for your records. After you finish reading, please click "next".

PROTOCOL DIRECTOR: Byron Reeves, 650 856-3644; reeves@stanford.edu, Department of Communication, Stanford University, Stanford, CA 94305-2050.

DESCRIPTION: We are doing research about how people use digital devices. The study will last for 2 weeks. Over the course of the 2 weeks, you will be
asked to fill out 2 surveys and install a research application on your phone (described below). You may also be offered some money to change your Internet usage behavior. The recruitment survey you have taken prior to arriving at this page is considered part of the study.

STUDY TASKS:

1. Install our research software onto your phone, and keep it running on the device. It runs passively, so you will not need to open it or interact with it after installation.
2. Take a 2-minute survey during the study period, for which you will be compensated $5. In this survey, you may be offered additional money to adjust your internet usage behavior, in which case you must choose to accept or decline.
3. Take a final 10-minute survey two weeks from now, after which you will receive $30 of compensation for participating for two weeks.

ABOUT OUR SOFTWARE: You will be asked to install software on your smartphone that will record the information on your screen every five seconds that your device is turned on for as long as the software remains on the device. The software will record anything that is on your screen including, for example, home screens, websites, photos, and messages that you send and receive. The software will also collect logs of all foreground application activity (i.e., the time when an app is opened and the name of the app); GPS (real time location) information every 5-10 minutes; and information about the current state of the phone's battery, memory, power, and functioning of the screenshot software itself. All of the data that we gather about your device usage will be kept strictly confidential. We will not use your name to identify your data file, and the data will be encrypted and stored on a secure Stanford server and storage system that will not be accessible to anyone outside of our research project. We will not share your data with anyone outside of the Stanford project. Identifiers might be removed from identifiable private information and, after such removal, the information could be used for future research studies or distributed to another investigator for future research studies without additional informed consent from you. You may also request the deletion of your data from our servers at any time by writing to reeves@stanford.edu. We will endeavor to complete your request within 30 days, unless otherwise required by law to retain the data. We may not be able to remove deidentified data that has been used in published researched or shared with other investigators as described above. We will delete screenshots from our servers after the research is complete. Additionally, we may disclose anonymized data, such as metrics derived from your phone activity (e.g., frequency of phone use, wordcount of on-screen text) or the types of places visited (e.g., restaurants, malls), to other academic researchers here at Stanford or other institutions for future, unrelated research studies, without additional informed consent from you. We may also ask you to complete one or more surveys as a part of this research.

DURATION: If you decide to participate in this study we ask you to participate for 2 weeks. You may discontinue the study at any time without penalty by removing the software from your device. You may also ask that your data be removed from our research database by contacting the Protocol Director (contact information above).

PAYMENT: You will be compensated $35 for participation: $30 for the keeping the software on your smartphone for two weeks, and taking a closing survey two weeks from now, plus $5 for filling out a 2-minute survey during the study period.

RISKS AND BENEFITS: Even though data will be stored on a secure Stanford server, there is a minimal risk of data breach and loss of confidentiality. Most of the data we collect will be analyzed by a computer, but a small portion of the screenshots may be viewed by Stanford researchers if computer analysis is not possible. Screenshots viewed by a researcher will be viewed only by trained Stanford researchers. We will not routinely review screenshots while they are collected and in most cases they will not be viewed at all by researchers. One possible benefit of participation in this research is that you may become more aware about your digital media use. However, we cannot and do not guarantee or promise that you will receive any benefits from this study. Based on information gained from this study, if the researchers may have serious concerns relating to matters such as severe depression, physical abuse, etc. or about your health and/or safety, the researchers may contact you and suggest a referral for your care.

SUBJECT'S RIGHTS: If you have read this form and have decided to participate in this project, please understand your participation is voluntary and you have the right to withdraw your consent or discontinue participation at any time without penalty or loss of benefits to which you are otherwise entitled. You have the right to refuse to answer particular questions. Your individual privacy will be maintained in all published and written data resulting from the study.

CERTIFICATE OF CONFIDENTIALITY: This research is covered by a Certificate of Confidentiality from the National Institutes of Health. This means that the researchers cannot release or use information, documents, or samples that may identify you in any action or suit unless you say it is okay. They also cannot provide them as evidence unless you have agreed. This protection includes federal, state, or local civil, criminal, administrative, legislative, or other proceedings. An example would be a court subpoena.

There are some important things that you need to know. The Certificate DOES NOT stop reporting that federal, state or local laws require. Some examples are laws that require reporting of child or elder abuse, some communicable diseases, and threats to harm yourself or others. The Certificate CANNOT BE USED to stop a sponsoring United States federal or state government agency from checking records or evaluating programs. The Certificate DOES NOT stop disclosures required by the federal Food and Drug Administration (FDA). The Certificate also DOES NOT prevent your information from being used for other research if allowed by federal regulations.

Researchers may release information about you when you say it is okay. For example, you may give them permission to release information to insurers, medical providers or any other persons not connected with the research. The Certificate of Confidentiality does not stop you from willingly releasing information about your involvement in this research. It also does not prevent you from having access to your own information.

CONTACT INFORMATION: *Questions, Concerns, or Complaints: If you have any questions, concerns, or complaints about this research study, its procedures, risks and benefits, you should ask the Protocol Director, Byron Reeves, 650 725-3033.
*Independent Contact: If you are not satisfied with how this study is being conducted, or if you have any concerns, complaints, or general questions about the research or your rights as a participant, please contact the Stanford Institutional Review Board (IRB) to speak to someone independent of the research team at (650)-723-2480 or toll free at 1-866-680-2906. You can also write to the Stanford IRB, Stanford University, 1705 El Camino Real, Palo

References

Albarracín, D., & Mitchell, A. (2004). The role of defensive confidence in preference for proattitudinal information: How believing that one is strong can sometimes be a defensive weakness. *Personality and Social Psychology Bulletin*, 30(12), 1565–1584. doi: 10.1177/0146167204271180

Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2020). The welfare effects of social media. *American Economic Review*, 110(3), 629-76.

Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, 6(14), eaay3539.

Althaus. (2002). American news exposure during times of national crisis. *Political Science & Politics.* 35, 03., 517–521.

Althaus & Tewksbury (2002). Agenda setting and the “new" news patterns of issue importance among readers of the paper and online versions of the New York Times. *Communication Research.* 29, 2. 180–207.

Arceneaux, K., Johnson, M., & Murphy, C. (2012). Polarized political communication, oppositional media hostility, and selective exposure. *The Journal of Politics*, *74*(1), 174-186.

Aridor, Guy, Drivers of Digital Attention: Evidence from a Social Media Experiment (March 29, 2022). *Available at SSRN*: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4069567.

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130-1132.

Baldassarri, D., & Bearman, P. (2007). Dynamics of political polarization. *American sociological review*, *72*(5), 784-811.

Banda, K. K., & Cluverius, J. (2018). Elite polarization, party extremity, and affective polarization. *Electoral Studies*, *56*, 90-101.

Barberá, P. (2014). How social media reduces mass political polarization. Evidence from Germany, Spain, and the US. *Job Market Paper, New York University*, *46*.

Barberá, P. (2020). Social Media, Echo Chambers, and Political Polarization. *Social Media and Democracy: The State of the Field, Prospects for Reform*, 34.

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber?. *Psychological science*, *26*(10), 1531-1542.

Baum & Groeling (2008). New media and the polarization of American political discourse. *Political Communication.* 25, 4. 345–365.

Bechmann, A., & Nielbo, K. L. (2018). Are we exposed to the same “news” in the news feed? An empirical analysis of filter bubbles as information similarity for Danish Facebook users. *Digital journalism*, *6*(8), 990-1002.

Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, 9(3), 226-232.

Bedeian, A. G., & Mossholder, K. W. (2000). On the use of the coefficient of variation as a measure of diversity. *Organizational Research Methods*, 3(3), 285-297.

Bennett. 2003. The burglar alarm that just keeps ringing: A response to Zaller. *Political Communication* 20, 2 (2003), 131–138.

Berkowitz, L., & Rogers, K. H. (1986). A priming effect analysis of media influences. *Perspectives on media effects*, *57*, 81.

Best, S. J., Chmielewski, B., & Krueger, B. S. (2005). Selective exposure to online foreign news during the conflict with Iraq. *Harvard International Journal of Press/Politics*, 10(4), 52-70.

Bimber, B. (2001). Information and political engagement in America: The search for effects of information technology at the individual level. *Political Research Quarterly*, 54(1), 53-67.

Bode, L. (2016). Political news in the news feed: Learning politics from social media. *Mass communication and society*, *19*(1), 24-48.

Boulianne, S. (2011). Stimulating or reinforcing political interest: Using panel data to examine reciprocal effects between news media and political interest. *Political Communication*, 28(2), 147-162.

Boutyline, A., & Willer, R. (2017). The social structure of political echo chambers: Variation in ideological homophily in online networks. *Political Psychology*, *38*(3), 551-569.

Brennan, R. L. (1992). Generalizability theory. *Educational Measurement: Issues and Practice*, *11*(4), 27-34.

Brians, C. L., & Wattenberg, M. P. (1996). Campaign issue knowledge and salience: Comparing reception from TV commercials, TV news and newspapers. *American Journal of Political Science*, 172-193.

Brinberg, M., Ram, N., Yang, X., Cho, M. J., Sundar, S. S., Robinson, T. N., & Reeves, B., 2021. The idiosyncrasies of everyday digital lives: Using the Human Screenome Project to study user behavior on smartphones. *Computers in Human Behavior*, 106570.

Bruns, A. (2017, September). Echo chamber? What echo chamber? Reviewing the evidence. In *6th Biennial Future of Journalism Conference (FOJ17)*.

Bryant, J., & Zillmann, D. (1984). Using television to alleviate boredom and stress: Selective exposure as a function of induced excitational states. *Journal of Broadcasting & Electronic Media*, 28(1), 1-20.

Cacciatore, M. A., Yeo, S. K., Scheufele, D. A., Xenos, M. A., Brossard, D., & Corley, E. A. (2018). Is Facebook making us dumber? Exploring social media use as a predictor of political knowledge. *Journalism & Mass Communication quarterly*, *95*(2), 404-424.

Carr, T. H., McCauley, C., Sperber, R. D., & Parmelee, C. M. (1982). Words, pictures, and priming: on semantic activation, conscious identification, and the automaticity of information processing. *Journal of Experimental Psychology: Human Perception and Performance, 8*(6), 757.

Chakraborty, Ghosh, Ganguly, & Gummadi. 2016. Dissemination Biases of Social Media Channels: On the Topical Coverage of Socially Shared News. In *ICWSM*. 559–562.

Chiatti, A., Yang, X., Brinberg, M., Cho, M. J., Gagneja, A., Ram, N., ... & Giles, C. L. (2017, December). Text extraction from smartphone screenshots to archive in situ media behavior. In *Proceedings of the Knowledge Capture Conference* (pp. 1-4).

Cho, M., Reeves, B., Ram, N., Robinson ,T. N., & Yang, X. 2022. Rhythms of smartphone use: Balancing media selections over time as a function of emotional valence, informational content, and time intervals of use. *Manuscript under review*.

Christensen, M. A., Bettencourt, L., Kaye, L., Moturu, S. T., Nguyen, K. T., Olgin, J. E., ... & Marcus, G. M. (2016). Direct measurements of smartphone screen-time: relationships with demographics and sleep. *PloS one*, 11(11), e0165331.

Cohen, J. (1992). A power primer. *Psychological bulletin*, *112*(1), 155.

Colleoni, E., Rozza, A., & Arvidsson, A. (2014). Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. *Journal of communication*, *64*(2), 317-332.

Conover, M. D., Ratkiewicz, J., Francisco, M. R., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political polarization on Twitter. *ICWSM*, *133*(26), 89-96.

Cutting, J. E., DeLong, J. E., & Nothelfer, C. E. (2010). Attention and the evolution of Hollywood film. *Psychological Science*, 21(3), 432-439.

Davies, J. C. (1965). The family's role in political socialization. *The Annals of the American Academy of Political and Social Science, 361*(1), 10-19.

Davis, N. T., & Dunaway, J. L. (2016). Party polarization, media choice, and mass partisan-ideological sorting. *Public Opinion Quarterly, 80*(S1), 272-297.

Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). Echo chambers: Emotional contagion and group polarization on facebook. *Scientific reports, 6*, 37825.

de Zúñiga, H. G., Borah, P., & Goyanes, M. (2021). How do people learn about politics when inadvertently exposed to news? Incidental news paradoxical direct and indirect effects on political knowledge. *Computers in Human Behavior, 121*, 106803.

Dhote, T., & Kumar, V. (2019). Long-duration storytelling: Study of factors influencing retention ability of brands. *Journal of Creative Communications*, 14(1), 31-53. doi:10.1177/0973258618822871

Dilliplane, S., Goldman, S. K., & Mutz, D. C. (2013). Televised exposure to politics: New measures for a fragmented media environment. *American Journal of Political Science, 57*(1), 236-248.

Dimmick, John, John Christian Feaster, and Gregory J. Hoplamazian. 2011. “News in the Interstices.” New Media & Society 13 (1):23–39.

Domke, D., Shah, D. V., & Wackman, D. B. (1998). Media priming effects: Accessibility, association, and activation. *International Journal of Public Opinion Research*, *10*(1), 51-74.

Druckman, J. N., & Levendusky, M. S. (2019). What do we measure when we measure affective polarization?. *Public Opinion Quarterly*, *83*(1), 114-122.

Dubois, E., & Blank, G. (2018). The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information, communication & society*, *21*(5), 729-745.

Dunst, L. (1993). Is it possible to get creative in 15 seconds? *Advertising Age*, 64(50), 18.

Eady, G., Nagler, J., Guess, A., Zilinsky, J., & Tucker, J. A. (2019). How many people live in political bubbles on social media? Evidence from linked survey and Twitter data. *Sage Open*, 9(1), 2158244019832705.

El-Din, D. M. (2016). Enhancement bag-of-words model for solving the challenges of sentiment analysis. *International Journal of Advanced Computer Science and Applications*, 7(1).

Elliot, A. J., & Devine, P. G. (1994). On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. Journal of Personality and Social Psychology, 67, 382-394.

Eveland Jr, W. P., & Seo, M. (2009). News and politics. In *Communication and social cognition* (pp. 305-330). Routledge.

Festinger, L. (1957). A theory of cognitive dissonance. Evanston, IL: Row, Peterson.

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*, *115*(27), E6106-E6115.

Fiske, S. T., & Russell, A. M. (2010). Cognitive processes. *The SAGE handbook of prejudice, stereotyping and discrimination*, 115-130.

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news exposure. *Public opinion quarterly*, *80*(S1), 298-320.

Freedman, P., & Goldstein, K. (1999). Measuring media exposure and the effects of negative campaign ads. *American journal of political Science*, 1189-1208.

Freedman, J. L., & Sears, D. O. (1965). Selective exposure. In *Advances in experimental social psychology* (Vol. 2, pp. 57-97). Academic Press.

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.

Friggeri, Adamic, Eckles, & Cheng (2014). Rumor Cascades. In *ICWSM.*

Galtung & Ruge (1965). The structure of foreign news: The presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers. *Journal of Peace Research.* 2, 1 (1965), 64–90.

Garrett, R. K. (2009). Echo chambers online?: Politically motivated selective exposure among Internet news users. *Journal of Computer-Mediated Communication*, *14*(2), 265-285.

Garrett, R. K. (2013). Selective exposure: New methods and new directions. Communication Methods and Measures, 7(3-4), 247-256.

Geiger, S. F., & Reeves, B., (1991). The effects of visual structure and content emphasis on the evaluation and memory for political candidates. Television and political advertising, 1, 125-143.

Geiger, S., & Reeves, B. (1993). The effects of scene changes and semantic relatedness on attention to television. *Communication Research*, 20(2), 155-175.

Gerber, A. S., Gimpel, J. G., Green, D. P., & Shaw, D. R. (2011). How large and long-lasting are the persuasive effects of televised campaign ads? Results from a randomized field experiment. *American Political Science Review*, 135-150.

Ghuman, A. S., Bar, M., Dobbins, I. G., & Schnyer, D. M. (2008). The effects of priming on frontal-temporal communication. *Proceedings of the National Academy of Sciences, 105*(24), 8405-8409.

Goldstein, D. G., McAfee, R. P., & Suri, S. (2011, June). The effects of exposure time on memory of display advertisements. *In Proceedings of the 12th ACM Conference on Electronic Commerce* (pp. 49-58). https://dl.acm.org/doi/pdf/10.1145/1993574.1993584

Guess, A. M. (2015). Measure for measure: An experimental test of online political media exposure. *Political Analysis*, 23(1), 59-75.

Guess, A., Nyhan, B., Lyons, B., & Reifler, J. (2018). Avoiding the echo chamber about echo chambers. *Knight Foundation*, *2*, 1-25.

Guess, A. M., Nyhan, B., & Reifler, J. (2020). Exposure to untrustworthy websites in the 2016 US election. *Nature Human Behavior*, *4*(5), 472-480.

Guess, A. M. (2021). (Almost) Everything in Moderation: New Evidence on Americans' Online Media Diets. *American Journal of Political Science*.

Hamaker, E (2012). Why researchers should think “within-person”: A paradigmatic rationale. *Handbook of Research Methods for Studying Daily Life* (The Guilford Press, New York), pp 43–61.

Hamilton, J. T. (2005). The market and the media. *Institutions of American Democracy: The Press*, 351-71.

Haim, M., Graefe, A., & Brosius, H. B. (2018). Burst of the filter bubble? Effects of personalization on the diversity of Google News. *Digital journalism*, *6*(3), 330-343.

Harcup & O’Neill (2001). What is news? Galtung and Ruge revisited. *Journalism Studies.* 2, 2 (2001), 261–280.

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychological bulletin*, *135*(4), 555.

Hermans, D., De Houwer, J., & Eelen, P. (2001). A time course analysis of the affective priming effect. *Cognition & Emotion*, *15*(2), 143-165.

Hsu, T. W., Niiya, Y., Thelwall, M., Ko, M., Knutson, B., & Tsai, J. L. (2021). Social media users produce more affect that supports cultural values, but are more influenced by affect that violates cultural values. *Journal of Personality and Social Psychology*.

Iyengar, Kinder, Peters, & Krosnick. (1984). The evening news and presidential evaluations. *Journal of Personality and Social Psychology.* 46, 4, 778.

Iyengar, S., Konitzer, T., & Tedin, K. (2018). The home as a political fortress: Family agreement in an era of polarization. *The Journal of Politics*, *80*(4), 1326-1338.

Jebb, A. T., Parrigon, S., & Woo, S. E. (2017). Exploratory data analysis as a foundation of inductive research. *Human Resource Management Review*, *27*(2), 265-276.

Jo, E., & Berkowitz, L. (1994). A priming effect analysis of media influences: An update.

Jungherr, A. (2018). Normalizing digital trace data. *Digital discussions. How big data informs political communication. Oxon: Routledge*, 19-45.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kaye, B. K., & Johnson, T. J. (2004). A Web for all reasons: uses and gratifications of Internet components for political content. *Telematics and informatics*, *21*(3), 197-223.

Kenski, K., & Stroud, N. J. (2006). Connections between Internet use and political efficacy, knowledge, and participation. *Journal of Broadcasting & Electronic Media*, 50(2), 173-192.

Kim, Y. (2011). The contribution of social network sites to exposure to political difference: The relationships among SNSs, online political messaging, and exposure to cross-cutting perspectives. *Computers in Human Behavior*, *27*(2), 971-977.

Kounios, J., Frymiare, J. L., Bowden, E. M., Fleck, J. I., Subramaniam, K., Parrish, T. B., & Jung-Beeman, M. (2006). The prepared mind: Neural activity prior to problem presentation predicts subsequent solution by sudden insight. *Psychological science*, *17*(10), 882-890.

Kramer, G. H. (1983). The ecological fallacy revisited: Aggregate-versus individual-level findings on economics and elections, and sociotropic voting. *American political science review*, *77*(1), 92-111.

Kwak, H., Lee, C., Park, H., & Moon, S. (2010, April). What is Twitter, a social network or a news media?. In *Proceedings of the 19th international conference on World wide web* (pp. 591-600).

Lang, A. (2009). The limited capacity model of motivated mediated message processing. In R. L. Nabi & M. B. Oliver (Eds.), T*he Sage handbook of media processes and effects* (pp. 99–112). Los Angeles, CA: Sage.

Lang, A., Zhou, S., Schwartz, N., Bolls, P. D., & Potter, R. F. (2000). The effects of edits on arousal, attention, and memory for television messages: When an edit is an edit can an edit be too much?. *Journal of Broadcasting & Electronic Media*, *44*(1), 94-109.

Lazer, D. (2020). Studying human attention on the Internet. *Proceedings of the National Academy of Sciences*, *117*(1), 21-22.

Lazer, D., Pentland, A. S., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., ... & Jebara, T. (2009). Life in the network: the coming age of computational social science. *Science (New York, NY)*, *323*(5915), 721.

Leighley, J. E., & Nagler, J. (1992). *Individual and systemic influences on turnout: Who votes?* 1984. The Journal of Politics, 54(3), 718-740.

Lelkes, Y. (2016). Mass polarization: Manifestations and measurements. *Public Opinion Quarterly*, *80*(S1), 392-410.

Lemke, J. L. (1998). Metamedia literacy: Transforming meanings and media. *Handbook of literacy and technology: Transformations in a post-typographic world*, *283301*.

Leung, L. (2020). Exploring the relationship between smartphone activities, flow experience, and boredom in free time. *Computers in Human Behavior*, *103*, 130-139.

Matthes, J., Nanz, A., Stubenvoll, M., & Heiss, R. (2020). Processing news on social media. The political incidental news exposure model (PINE). *Journalism*, *21*(8), 1031-1048.

McBirnie, Abigail. 2008. "Seeking Serendipity." *Aslib Proceedings* 60 (6):600–618.

McLeod, D. M., Kosicki, G. M., & McLeod, J. M. (2009). Political communication effects. In *Media Effects* (pp. 244-267). Routledge.

Messing & Westwood (2014). Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online. *Communication Research.* 41, 8, 1042–1063.

Miller, N., & Campbell, D. T. (1959). Recency and primacy in persuasion as a function of the timing of speeches and measurements. *The Journal of Abnormal and Social Psychology, 59*(1), 1.

Mohammad, S. (2018, July). Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 174-184).

Molenaar, P. C. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika*, *50*(2), 181-202.

Molenaar, P. C. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, *2*(4), 201-218.

Molenaar, P. C. (2008). On the implications of the classical ergodic theorems: Analysis of developmental processes has to focus on intra-individual variation. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *50*(1), 60-69.

Molenaar, P. C., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current directions in psychological science*, *18*(2), 112-117.

Molotch & Lester (1974). News as purposive behavior: On the strategic use of routine events, accidents, and scandals. *American Sociological Review*. 101–112.

Molyneux, L. (2018). Mobile news exposure: A habit of snacking. *Digital Journalism*, *6*(5), 634-650.

Mondak. (1995). Newspapers and political awareness. *American Journal of Political Science*, 513–527.

Morales, A. J., Borondo, J., Losada, J. C., & Benito, R. M. (2015). Measuring political polarization: Twitter shows the two sides of Venezuela. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *25*(3), 033114.

Muise, D., Hosseinmardi, H., Howland, B., Rothschild, D., Mobius, M., Watts, D. (2022). Quantifying Partisan News Diets on Television and Online. *Manuscript under review*. Preprint link for dissertation committee.

Muise, D., Lu, Y., Pan, J., & Reeves, B. (2022). Selectively localized: Temporal and visual structure of smartphone screen activity across media environments. *Mobile Media & Communication*, 20501579221080333.

Muise, D., Reeves, B., and Pan, J. (2017). "What is News?" Realigning the News Definition with Millions of Consumer Screenshots." *Computation + Journalism*. Northwestern University.

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*.

Nithyanand, R., Schaffner, B., & Gill, P. (2017). Measuring offensive speech in online political discourse. In *7th USENIX workshop on free and open communications on the internet (FOCI 17)*.

Nguyen, C. T. (2020). Echo chambers and epistemic bubbles. *Episteme*, *17*(2), 141-161.

Oeldorf-Hirsch, A. (2018). The role of engagement in learning from active and incidental news exposure on social media. *Mass Communication and Society*, *21*(2), 225-247.

Ooms, Jeroen (2020). *hunspell*: High-Performance Stemmer, Tokenizer, and Spell Checker. *R package version 3.0.1*. https://CRAN.R-project.org/package=hunspell

Oulasvirta, A., Rattenbury, T., Ma, L., & Raita, E. (2012). Habits make smartphone use more pervasive. *Personal and Ubiquitous computing*, 16(1), 105-114.

Oulasvirta, A., Tamminen, S., Roto, V., & Kuorelahti, J., (2005), April. Interaction in 4-second 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.

Paisley, W. (1984). Communication in the communication sciences. *Progress in communication sciences*, *5*, 1-43.

Pariser, E. (2011). *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin.

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,* 1-20.

Peng, T. Q., & Zhu, J. J., (2020). Mobile Phone Use as Sequential Processes: From Discrete Behaviors to Sessions of Behaviors and Trajectories of Sessions. *Journal of Computer-Mediated Communication*, 25(2), 129-146.

Pennington, J., Socher, R., & Manning, C. D. (2014, October). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532-1543).

Peterson, E., Goel, S., & Iyengar, S. (2021). Partisan selective exposure in online news exposure: Evidence from the 2016 presidential campaign. *Political science research and methods*, *9*(2), 242-258.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In *Communication and persuasion* (pp. 1-24). Springer, New York, NY.

Piantadosi, S., Byar, D. P., & Green, S. B. (1988). The ecological fallacy. *American journal of epidemiology*, 127(5), 893-904.

Price, V., Ritchie, L. D., & Eulau, H. (1991). Cross-level challenges for communication research: Epilogue. *Communication Research*, *18*(2), 262-271.

Prior (2003). Any good news in soft news? The impact of soft news preference on political knowledge. *Political Communication.* 20, 2 (2003), 149–171.

Prior, M. (2009). Improving media effects research through better measurement of news exposure. *The Journal of Politics*, 71(3), 893-908.

Prior, M. (2010). You've either got it or you don't? The stability of political interest over the life cycle. *The Journal of Politics*, 72(3), 747-766.

Prior, M. (2013). Media and political polarization. *Annual Review of Political Science, 16*, 101-127.

Reeves, B. (1989). Theories about news and theories about cognition: Arguments for a more radical separation. *American Behavioral Scientist*, *33*(2), 191-198.

Quattrone, G., & Tversky, A. (1988). Contrasting rational and psychological analyses of political choice. *American Political Science Review*, 82, 719–736.

Reeves, B., and N. Ram, T.N. Robinson, J. J. Cummings, L. Giles, J. Pan, A. Chiatti, MJ Cho, K. Roehrick, X. Yang, A. Gagneja, M. Brinberg, D. Muise, Y. Lu, M. Luo, A. Fitzgerald & L. Yeykelis, 2019. *Human Computer Interaction*, 1-52.

Reeves, B., T.N. Robinson & N. Ram, (2020). Time for the Human Screenome Project. *Nature*, January 15, vol 577: 314-317.

Reeves, B., Thorson, E., Rothschild, M. L., McDonald, D., Hirsch, J., & Goldstein, R., 1985. Attention to television: Intrastimulus effects of movement and scene changes on alpha variation over time. *International Journal of Neuroscience*, 27(3-4), 241-255.

Rogers, S. C. (1995). How to create advertising that works. *Journal of Business & Industrial Marketing,* 10(2), 20–33.

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.

Roskos-Ewoldsen, D. R., Klinger, M. R., & Roskos-Ewoldsen, B. (2007). Media priming: A meta-analysis. *Mass media effects research: Advances through meta-analysis*, 53-80.

Rozner, G. (2018). Inside the intellectual dark web. *Institute of Public Affairs Review: A Quarterly Review of Politics and Public Affairs, The*, 70(3), 6-11.

Rubin, Victoria L., Jacquelyn Burkell, and Anabel Quan-Haase. 2011. "Facets of Serendipity in Everyday Chance Encounters." *Information Research* 16 (3):27–27.

Rudkowsky, E., Haselmayer, M., Wastian, M., Jenny, M., Emrich, Š., & Sedlmair, M. (2018). More than bags of words: Sentiment analysis with word embeddings. *Communication Methods and Measures*, *12*(2-3), 140-157.

Rula, J. P., Jun, B., & Bustamante, F. (2015, February). Mobile AD (D) Estimating Mobile App Session Times for Better Ads. In *Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications* (pp. 123-128).

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161.

Salvatore, S., & Valsiner, J. (2010). Between the general and the unique: Overcoming the nomothetic versus idiographic opposition. *Theory & Psychology*, 20(6), 817-833.

Scheufele, D. A. (2000). Agenda-setting, priming, and framing revisited: Another look at cognitive effects of political communication. *Mass communication & society*, *3*(2-3), 297-316.

Sears, D. O., & Freedman, J. L. (1967). Selective exposure to information: A critical review. *Public Opinion Quarterly*, *31*(2), 194-213.

Shah, Watts, Domke, Fan, Fibison. (1999). News coverage, economic cues, and the public's presidential preferences, 1984-1996. *The Journal of Politics* 61, 4 (1999), 914–943.

Shim, H., You, K. H., Lee, J. K., & Go, E. (2015). Why do people access news with mobile devices? Exploring the role of suitability perception and motives on mobile news use. *Telematics and Informatics*, *32*(1), 108-117.

Singh, S. N., Rothschild, M. L., & Churchill, G. A. (1988). Recognition versus recall as measures of television commercial forgetting. *Journal of Marketing Research*, 25(1), 72.

Soroka, S. N. (2006). Good news and bad news: Asymmetric responses to economic information. *Journal of Politics*, 68, 372–385.

Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *Journal of management information systems*, *29*(4), 217-248.

Stoker, G., Hay, C., & Barr, M. (2016). Fast thinking: Implications for democratic politics. *European Journal of Political Research*, *55*(1), 3-21.

Stroud, N. J. (2010). Polarization and partisan selective exposure. *Journal of Communication*, *60*(3), 556-576.

Tewksbury, D., Weaver, A. J., & Maddex, B. D. (2001). Accidentally informed: Incidental news exposure on the World Wide Web. *Journalism & Mass Communication Quarterly*, *78*(3), 533-554.

Thorson, K., Cotter, K., Medeiros, M., & Pak, C. (2021). Algorithmic inference, political interest, and exposure to news and politics on Facebook. *Information, Communication & Society*, *24*(2), 183-200.

Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of personality and social psychology*, *90*(2), 288.

van Aelst, P., & De Swert, K. (2009). Politics in the news: Do campaigns matter? A comparison of political news during election periods and routine periods in Flanders (Belgium). *Communications*, *34*(2), 149-168.

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.

van Damme, K., Martens, M., Van Leuven, S., Vanden Abeele, M., & De Marez, L. (2020). Mapping the mobile DNA of news. Understanding incidental and serendipitous mobile news exposure. *Digital Journalism*, *8*(1), 49-68.

van Erkel, P. F., & Van Aelst, P. (2020). Why don't we learn from social media? Studying effects of and mechanisms behind social media news use on general surveillance political knowledge. *Political Communication*, 1-19.

van Orden, G.C., Holden, J.G., Turvey, M.T. (2003). Self-organization of cognitive performance. *Journal of Experimental Psychology: General*, 152, 331–350.

Weaver, A. J. (2011). The role of actors' race in white audiences' selective exposure to movies. *Journal of Communication*, 61(2), 369-385.

Weeks, B. E., Lane, D. S., Kim, D. H., Lee, S. S., & Kwak, N. (2017). Incidental exposure, selective exposure, and political content sharing: Integrating online exposure patterns and expression on social media. *Journal of Computer-Mediated Communication*, *22*(6), 363-379.

Weiss GM (2004) Mining with rarity: a unifying framework. *SIDKDD Explorations* 6(1):7–19

West, E. A., & Iyengar, S. (2020). Partisanship as a social identity: Implications for polarization. *Political Behavior*, 1-32.

White (1950). The "gate keeper": A case study in the selection of news. Journalism & Mass Communication Quarterly 27, 4, 383–390.

Winter, S., Metzger, M. J., & Flanagin, A. J. (2016). Selective use of news cues: A multiple-motive perspective on information selection in social media environments. *Journal of Communication*, *66*(4), 669-693.

Xu, S., & Lorber, M. F. (2014). Interrater agreement statistics with skewed data: Evaluation of alternatives to Cohen's kappa. *Journal of Consulting and Clinical Psychology*, 82(6), 1219.

Yang, T., Majo-Vazquez, S., Nielsen, R. K., & González-Bailón, S. (2020). Exposure to News Grows Less Fragmented with Increase in Mobile Access. *Available at SSRN* 3564826.

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

Yarkoni, T. (2022). The generalizability crisis. *Behavioral and Brain Sciences*, *45*.

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.

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.