The Search for Real-World Media Effects

Thomas J. Leeper[i]
London School of Economics and Political Science

Prepared for Inclusion in Zoe, Oxley, and David Redlawsk, eds., *Oxford Encyclopedia of Political Decision Making*, Oxford: Oxford University Press.

Summary
The study of media effects entails the empirical association of two things: measures of media content or experience and measures of audience outcomes. Any quantitative evidence of correlation between the two—in tandem with assumptions about causal ordering and an absence of spuriousness—constitutes evidence of media effects. At stake in the search are three challenges: the measurement of the outcomes, the measurement of the media and individuals' exposure to it, and the tools and techniques for associating the two.

While measuring the outcomes potentially affected by media exposure is in many ways trivial (surveys, election outcomes, and online behavior provide numerous measurement devices), the latter two aspects of studying media effects present nearly insurmountable empirical difficulties short of ambitious experimental design. Despite these challenges, media effects research has been preoccupied for much of its history by an effort to develop and apply survey-based measures of individual media exposure that serve as the empirical basis for studying media effects. The effort to use such measures to generate causal insight into media effects has ultimately distracted from the design of both causally credible methods such as field experiments and thicker descriptive research on the content and experience of media. Outside of the laboratory, we understand media effects too little despite considerable time and effort.

The canonical approach for assessing such effects: namely, using survey questions about individual media experiences to measure the putatively causal variable and correlating those measures with other measured outcomes suffers from substantial limitations. Experimental—and sometimes quasi-experimental—methods provide definitely superior causal inference about media effects and provide a uniquely fruitful path forward for insight into media and their effects.

Keywords: media effects, media, causal inference, experiments, field experiments, quasi-experiments, surveys, media exposure

Introduction

The study of media effects entails the empirical association of two things: measures of media content or experience and measures of audience outcomes. Any quantitative evidence of correlation between the two—in tandem with assumptions about causal ordering and an absence of spuriousness—constitutes evidence of media effects. Social scientists are particularly interested in any such effects on the public's perceptions of the social and political world, their knowledge or lack thereof about the same, their preferences over goods, candidates, or issues, and finally their behavior. The search for media effects takes many forms and this chapter focuses on the search for those effects outside the confines of experimental laboratories, in the buzzing, blooming confusion of everyday life.

At stake in the search are three challenges: the measurement of the outcomes, the measurement of the media and individuals' exposure to it, and the tools and techniques for associating the two. While measuring the outcomes potentially affected by media exposure is in many ways trivial (surveys, election outcomes, and online behavior provide numerous measurement devices), the latter two aspects of studying media effects present nearly insurmountable empirical difficulties short of ambitious experimental design. Despite these challenges, media effects research has been preoccupied for much of its history by an effort to develop and apply survey-based measures of individual media exposure that serve as the empirical basis for studying media effects. Despite Prior's (2013) call to arms that "developing better measures of media exposure is a pressing goal" (621), the effort to do so has been a largely failed exercise that has left social scientists with little credible insight into media effects outside of laboratory settings—precisely those locations where such effects matter the most. The effort to use survey-based measures to generate causal insight into media effects has ultimately distracted from the design of both causally credible methods such as field experiments and thicker descriptive research on the content and experience of media. Outside of the laboratory, we understand media effects too little despite considerable time and effort.

Laboratory experiments have demonstrated causal possibilities, but can generalize weakly given the self-selected nature of media experiences (Gaines and Kuklinski 2011; Arceneaux and Johnson 2012; Leeper 2017) and the arbitrary selection of treatments, outcomes, and samples in much experimental work (Druckman and Leeper 2012b). Field experimental studies therefore present the best path forward for insights into media effects outside of such settings because of their causal credibility and the advantage of true experiments—relative to so-called "natural experiments" (Sekhon and Titiunik 2012)—at offering insight into anything beyond quirks of causality. But just as field experiments present an ideal path for obtaining credible and realistic insights into media effects, thick descriptive methods spanning the qualitative-quantitative divide present promising opportunities for studying media content and media experiences that are likely to generate far more useful insights than thin-descriptive survey measures of media exposure. Like Graber's (1988) seminal use of in-depth interviews, methods that go beyond "mere exposure" are vital for understanding the complexities of media experiences that might be the basis for "media effects".

What follows here is a discussion the concept of "media effects" and the evidentiary standards necessary to establish that media have a causal effect on politically relevant outcomes. I then discuss the substantial limitations in the canonical approach for assessing such effects: namely, using survey questions about individual media experiences to measure the putatively causal variable and correlating those measures with other measured outcomes. Next, I demonstrate how experimental—and sometimes quasi-experimental—methods provide definitely superior causal inference about media effects and conclude with a discussion of how these methods—and others—might be fruitfully deployed moving forward.

What are Media Effects and How Would We Know Them When We See Them?

Like any causal relationship, there are two ways to frame the question of media effects: either as a "backward causal question" emphasizing the role media variables—relative to many others—might play in the production of observed outcomes, or as a "forward causal question" emphasizing how outcomes might differ across counterfactual values of media variables (Gelman and Imbens 2013).[ii] The backward-looking approach takes outcomes as phenomena to be explained and seeks out explanations for what might have caused them, ultimately attempting to assess the absolute or relative size of the media contribution to those outcomes. These outcomes might be macro level like election outcomes or public discourse or they might be micro level outcomes like individual beliefs, opinions, affect, cognition, physiology, or behaviors (Potter 2011). Less abstractly: Why did individuals vote for Donald Trump in the 2016 US Presidential election? Why did Britain vote leave in the 2016 referendum on European Union membership? "Media," however defined and conceptualized, might be sought out as one among many possible causes of these outcomes. Media variables typically taking the form of metrics of media content or metrics of individuals' exposure to, attention to, or reception of said content.

The forward-looking approach instead takes media or a feature thereof to be a well-defined and perhaps *manipulable* variable that generates different outcomes across counterfactual values of the variable (Rubin 1978; Holland 1986). The outcomes of interest are the same but the forward-looking approach attempts to reduce "media" as a concept to an isolatable event, experience, or exposure and assess how realized outcomes compare to counterfactual outcomes where "media" were different. Less abstractly: What if the Hillary Clinton campaign had spent more on television advertising in swing states in the 2016 Presidential election, would vote shares have been different? What if media had covered the

Leave campaign's '£350 million per week for the NHS' claim differently, would vote intentions in the 2016 referendum have been different?

The phenomena and the causal relationships are the same, but the backward and forward framings of media effects steer attention toward specific kinds of questions and specific kinds of research designs. In the backward-looking framing, research in search of media effects substantiates effects when variation in outcomes across variations in media variables persists once other explanations for that outcome variation have been considered and controlled for. In the forward-looking framing, research in search of media effects substantiates effects when variation in outcomes manifests in response to a real or approximate manipulation of a media variable. Backward causal questions are exploratory; forward causal questions are experiment-like. Whereas forward-looking questions generate a definitive statement about the direction(s) and size(s) of media influence on outcomes of interest, backward-looking questions lead only to further questions or new hypotheses.

Humans tend to think about causal effects – including those of media – in backward-looking terms, so we naturally gravitate toward trying to answer those questions directly. But as Gelman and Imbens (2013) argue, these questions never lead to clear answers because they ultimately generate correlational evidence influenced by unobserved or unobservable additional factors. In Hovland's (1959) words: "while the conceptualization of the survey researcher is often very valuable, his [sic] correlational research design leaves much to be desired" (15). Instead, to understand media effects researchers need to transform backward-looking questions into forward-looking questions or treat the answers to backward-looking questions as exploratory steps that lead toward new forward-looking questions. A backward-looking question at best generates hypotheses about possible causes but does not rule out causes nor definitively clarify the magnitude of causal effects because there can always be some other set unobserved factors that explain away any observed patterns, or mask causality under apparent non-correlation. The researcher must identify and measure not just the media variable but all other potential causes.[iii]

Despite the difficulty of answering backward-looking causal questions about media effects, researchers continue to search for explanations of outcomes by associating those outcomes with measures of media variables controlling for a subset of other observable, measurable phenomena. Such analysis is facilitated by the frequent inclusion of coarse, unreliable, and error-prone survey-based measures of media exposure in nearly all election surveys and many public opinion polls. Despite the general, philosophical challenges to backward causal inference, much research continues to proceed from an assumption that not only is backward-looking causal inference possible but also that survey-based measures have any utility at all in causal inference. In an infamous example, Bartels (1993) regresses various election-related, individual-level political outcomes on two American National Election Studies items measuring television viewing and daily newspaper readership, controlling for party identification, age, education, and race. The results suggest larger than anticipated effects due to corrections for measurement error in the media measures providing an apparently substantial corrective to a field the author terms "one of the most notable embarrassments of modern social science" (267).

Yet to answer forward-looking questions requires conceptualizations of and measures of media experiences that do not in any way resemble the survey-based measures in wide use in the late 20th and early 21st centuries. Rather than asking "how much did media matter in the last US Presidential campaign?" and attempting to answer that question by regressing vote choice on a self-reported, survey-based measure of media exposure and some possible confounds, a more credible forward-looking approach attempts narrowly to understand whether and to what degree an isolatable media experience—such as viewing a debate, seeing a television advertisement, reading a particular news story—affected individual vote choice or aggregate election outcomes. Doing so requires both narrowness in research question but also attention to measurement of a specific event rather than abstract media experiences (Prior 2007). For example, Fridkin et al. (2007) used randomized exposure to a presidential debate to understand what impact the debate had on a variety of outcomes. Rather than try to define and measure campaign effects broadly, they focus on an isolatable experience. Similarly, Albertson and Lawrence (2009) randomly *encourage* viewing of an educational television program to understand the effect of this specific event – rather than some abstract definition of television generally – on knowledge and attitudes. Before explaining why survey-based measures of media experience are particularly flawed for understanding media effects, it is important to see how these kinds of experimental approaches provide a straightforward design for answering forward causal questions but no method provides a straightforward design for answering backward causal questions about media effects.

Design Trumps Analysis in Studying Media Effects

A causal effect (of media) for an individual is conventionally understood as a difference between two or more *potential* outcomes that this individual might have expressed had they been exposed to varying values of a media variable (see Rosenbaum and Rubin 1982; Holland 1986; Gerber and Green 2012). To take a canonical example, an individual's opinion on whether to tolerate a rally by a hate group might be affected by different media portrayals of the issue ("framing"), such as coverage that emphasizes free speech considerations versus coverage that emphasizes public safety considerations (Nelson et al. 1997). The individual-level causal effect, *TE*, is defined as the difference in the individual's, *i*, potential opinion, Y_{it}, when they are exposed to one or the other framing at a given point in time, *t*:

$$TE_{it} = Y_{\{it,Free\ Speech\}} - Y_{\{it,Public\ Safety\}}$$

Defined in this way, these two media experiences constitute the complete set of possible media experiences available at time *t*.[iv] Because individuals can only experience one of these forms of media content,[v] the individual-level media effect is unobservable due to the fundamental problem of causal inference (Holland 1986). We can never know how media affect an individual without insight into unobservable counterfactuals where their contemporaneous media experiences are different.[vi]

Causal inference, however, often proceeds from the idea that while individual-level effects are not observable, an *average treatment effect* across a population (or sample thereof) is observable and provides useful insight into the central tendency of the individual-level effect distribution. Because of the following equality:

$$\begin{aligned} ATE &= E(TE_t) \\ &= E\left(Y_{\{t,Free\ Speech\}} - Y_{\{t,Public\ Safety\}}\right) \\ &= E\left(Y_{\{t,Free\ Speech\}}\right) - E\left(Y_{\{t,Public\ Safety\}}\right) \end{aligned}$$

we can reach inferences about the average effect of the free speech framing (relative to public safety framing) by comparing average outcomes among individuals exposed to each type of content, provided we are willing to assume that these individuals receive content independent of the values they would take for $Y_{Free\ Speech}$ and $Y_{Public\ Safety}$. In an experimental setting we can assume this independence by design because physical randomization of individuals to experiences operates without regard for each individual's schedule of potential outcomes. In all other (observational) research designs, we must arrive upon that independence only *conditionally* by mathematically conditioning on factors that are theorized to influence both individuals' exposure to particular media at a particular point in time and their potential outcomes. Fully identifying and measuring these other factors is daunting.

This essential difference between how we draw causal inferences in experimental and observational research—in the former case by design and in the latter case only by careful and complete measurement of confounding factors—highlights why experimental studies are seen to provide a "gold standard." However, this particularly favorable standing among alternative research designs is not absolute. Indeed, experiments taking place in laboratory settings – including some of the earliest experimental research in media effects by Iyengar and Kinder (1987) – are seen as particularly limited in value. Yet as Hovland (1959) argued decades ago, integrating observational (survey) and (laboratory) experimental methods "will require on the part of the experimentalist an awareness of the narrowness of the laboratory in interpreting the larger and more comprehensive effects of communication. It will require on the part of the survey researcher a greater awareness of the limitations of the correlational method as a basis for establishing causal relationships" (14). Experimental designs for studying media effects provide the possibility of clear inference about average media effects but they do not necessarily do so in a way that is consistently useful and the reliance upon experimental manipulation limits the degree to which experimentation *explains* mediatized phenomena. Operating outside of the laboratory and beyond the scope of survey-based measures of media exposure is likely to be particularly fruitful and is the focus of this chapter.

Strictly speaking, experiments (be they in a laboratory, survey or field setting) provide insights into causal possibilities. Media effects experiments test whether a particular media variation *can* cause an outcome, within the implicit constraints of the sample, setting, and used in the experiment (see Shadish, Cook, and Campbell 2001, esp. ch. 13). Evidence that a given experience is, on average, effectual in a particular time and place for a particular set of individuals does not mean that the same results would be obtained elsewhere. Experimental evidence of media effects must always be read as "media *can* cause" not "media *do* cause." But experiments are also limited by the prospective, forward-looking nature of the research design. Short of massive-scale, field-based interventions into everyday life, experiments also cannot generate inferences of the form “media *did* cause.” For example, Feezell (2017) demonstrates that social media can serve an agenda setting function but not that they have in any particular instance outside of the experimental context. Druckman et al. (2017) show that mediatized messages can further spread through interpersonal discussions but not that they have in any particular instance outside of the experimental context. Searles et al. (2017) show that male and female voices can be differentially effective in campaign advertising but not that they are in any particular instance outside of the experimental context. Extrapolation beyond the experimental setting, sample, treatment, and outcome measures requires assumptions about or an explicit model for the transportability of the causal effects. This means experiments are typically powerless on their own to provide retrospective or historical insight and thus powerless to answer the kinds of backward-looking causal questions that social scientists frequently gravitate toward.

This is particularly worth in-depth consideration given that media experiences, the effects of which researchers might desire to know, are not commonly randomly assigned (Hovland 1959; Bennett and Iyengar 2008; Arceneaux and Johnson 2012; Leeper 2017) nor do they consist of strictly captive exposure to forced stimuli (Druckman, Fein, and Leeper 2012). Media content and audience exposure to that content is anything but random.[vii] Experiments thus provide “gold standard” causal insight into experiences but only to the extent that the variation introduced by experimental control resembles the real-world variation in media experiences that researchers might desire to understand and that such experiences are prone to be easily randomized.

On face value then observational methods of obtaining causal inference about media effects would seem to have some advantages over these narrow experimental approaches. For example, observational methods would allow a greater flexibility over the sample of individuals, settings, causes, and outcomes being studied given that the experiment-eligible populations of individuals, settings, causes, and outcomes are a non-representative subset of this hyperpopulation of interest. Similarly, observational methods may be deployed in service to retrospective questions that are impossible for prospective experimental techniques to answer. And observational methods rely upon naturalistic – rather than researcher-forced – variation in media experiences, minimizing concerns about the artifice of the experimental experience. But these apparent superiorities of observational approaches are frequently illusory. Single-study characteristics that imply generalizability, such as representative sampling of causes, outcomes, settings, and units are only useful for learning about media effects to the extent those ostensibly more "general" research designs also offer causal identification. Frequently, they do not. Thus the oft-mentioned trade-off is not between internal validity and external validity but between clear identification of a possible causal effect and an alternative design that offers neither clarity about internal nor external validity.

An increasingly popular middle ground between experimental media effects research and observational media effects research are so-called "quasi-experiments" or "natural experiments" (Shadish, Cook, and Campbell 2001). Unlike experiments, quasi-experiments do not involve the active intervention of the researcher but instead analysis of variation in outcomes across random or as-if-random interventions generated by other forces (such temporal and geographical discontinuities, the random spread of technologies, weather patterns or geological factors, or lotteries administered for other purposes). Such designs might attempt to understand the direct effect of a randomized media phenomenon or use randomization-like variation in something else to instrument for a media coverage, access, or exposure. For example, researchers have studied how electoral outcomes vary geographically across areas affected early or late by the non-random but also not wholly systematic rollout of cable television, broadband internet (Lelkes, Sood, and Iyengar 2015) or Fox News (DellaVigna and Kaplan 2007; Clinton and Enamorado 2014). Other quasi-experimental approaches to media effects use discontinuities in radio or television signal strength in Russia (Enikolopov, Petrova, and Zhuravskaya 2011) or Silvio Berlusconi's Mediaset network in Italy (Durante, Pinotti, and Tesei N.d.), arbitrary channel positioning of Fox News across US cable providers (Martin and Yurukoglu 2017), or the unintentional overlap of US competitive-state media markets into neighboring non-competitive districts (Huber and Arceneaux 2007; Krasno and Green 2008).

Relying on strong assumptions about the randomness of these "natural" interventions, quasi-experiments provide an observational research design that generates more credible causal inference than traditional correlational designs given that areas affected and unaffected by such interventions are considered to be similar with the intervention occurring as-if-random (for methodological discussion, see Sovey and Green 2011; Keele and Titiunik 2014). The advantage of quasi-experiments over researcher-administered experiments is the ability to gain retrospective and historical insight into outcomes that might have been impacted by quasi-experimental intervention (see, for example, Voigtlaender and Voth 2014). Even more so than experiments, however, they have a narrowness of scope – and thus a severe inferential localness – due to the rarity of naturally occurring randomization. True lotteries occur (e.g., with the US Vietnam War draft lottery [Angrist 1990; Erikson and Stoker 2011]) but most quasi-experiments leverage a strong assumption of randomness applied to a one-off media occurrence. Rather than being a middle ground between poorly identified observational methods and well-identified experimental methods, quasi-experiments carry the strengths and limitations of both approaches, such that they serve as very useful in the rare instances in which they occur as they provide a retrospective approach to answering a forward-looking causal question but may tend to cumulate poorly given the typical impossibility of replicating a given non-experimental intervention. But as we have seen, all causally oriented research suffers from some degree of localness.

Despite the promise, quasi-experiments provide useful but ultimately narrow historical insight into media effects. Experimental methods offer a superior alternative given the inherent repeatability of interventions (although not necessarily the settings in which they are randomized). Whereas both experiments and quasi-experiments offer a precise definition of a cause and precise statement of effects; observational methods for measuring media effects suffer from the lack of such precision because – despite decades of scholarly effort – we often cannot effectively conceptualize let alone measure variations in individual media experiences. If we do not know how to ensure that two media experiences differ in one and only one way, observational research leaves us with a "bundles of sticks" problem (Sen and Wasow 2016) where we attribute differences in outcomes to "media" (vaguely defined and coarsely measured) without any clear insight into what part of "media" is producing effects. Observational methods for obtaining causal inference only achieve credibility when we can define, measure, and control for all other differences in experiences; this is something we cannot do. In essence, we learn nothing.

Measuring Media and Media Exposure

Such pessimism about observational approaches is warranted. While much social science research would seem to imply that researchers believe it is possible to measure media exposure using surveys, this is simply not the case. Though numerous scholars have advocated for improved measures of media exposure and attempted to expose the deficiencies or advantages of particular approaches (Althaus and Tewksbury 2007; Dilliplane, Goldman, and Mutz 2012; Goldman, Mutz, and Dilliplane 2013; Tewksbury, Althaus, and Hibbing 2011; Prior 2009a,b, 2013; Guess 2014; Jerit et al. 2016; Bartels 2003; Chaffee and Schleuder 1986; Freedman and Goldstein 1999; de Vreese and Boomgaarden 2006; Prior 2003; Dilliplane 2011; Price and Zaller 1993; Eveland, Seo, and Marton 2002; Eveland and

Scheufele 2000; Eveland et al. 2005; Eveland, Hutchens, and Shen 2009; Slater, Goodall, and Hayes 2009; Garrett, Carnahan, and Lynch 2013), this collective effort at obtaining complete measures of media exposure is fundamentally flawed. This goes beyond the use of such measures in causal inference.

Consider, for example, a few common ways of measuring media exposure using survey self-reports. We might ask individuals to report whether they have been exposed to (or attentive to) a particular source, a particular medium, or a particular event. Alternatively, we might ask for a ranking of the degree of attention to certain sources (like CNN or Fox News) or news content (for example about a particular piece of legislation or world event). Alternatively, we might ask for degree of attentiveness or ratings of intensity of use of various media (like television or internet news). Alternatively, we might ask for time-based or frequency-based measures that count the number of days, hours, or minutes spent with media. All of these approaches might vary in their source-specificity from an abstract medium (e.g., television, newspapers) to specific sources (e.g., World News Tonight on ABC), and vary in their content-specificity from abstract topics (e.g., news about politics and international affairs) to specific facts (e.g., news of Donald Trump's alleged affair with pornographic actress Stormy Daniels during the pregnancy of his third wife, Melania Trump). And each can vary in the granularity of time used to measure such exposure or to rank exposure to media alternatives: we might ask about typical behavior, behavior the previous week, behavior that day, or even hours, minutes, or seconds of time use. These measures tell us what people believe their media experiences are, but only coarsely and with substantial error. The public substantially over-report media attention for a variety of reasons (Prior 2013) and efforts at improved measures (like those just alluded to) have not produced any degree of consensus on how to improve media use measures.

The challenges discussed in the literature are, in my view, quite superficial. These tend to include measurement error, over-reporting, lack of over-time reliability, and social desirability biases. But if one wants to understand *media effects* apart from understanding *media use*, then a larger epistemological issue is whether it makes sense to talk about effects of ill-defined causes. If we imbue an error-prone, biased survey-based measure of media attention with causal meaning, what kinds of claims can we generate? Consider, for example, a claim by Kull, Ramsay, and Lewis (2003) in their study of misperceptions related to the Iraq War that in a regression analysis of misperceptions controlling for demographics, "the respondent's primary source of news is still a strong and significant factor; indeed, it was one of the most powerful factors predicting misperception" (587). In other words, Fox News had the effect of leading the public to be misinformed about the Iraq War. Setting aside the low credibility of causal ordering in this design,[viii] what does the "Fox News effect" mean here?

Rubin (1990) describes how causal inference in the potential outcomes framework requires a stable unit treatment value assumption (SUTVA). While most research focuses on the non-interference part of this assumption, SUTVA also requires that there is only version or form of the treatment (see Sinclair 2012). This *treatment homogeneity* is perhaps the most overlooked assumption of causal inference. If a person is characterized as having an identical value of treatment (for example *FoxViewer=1*) then the assumption is that this person's treatment is the same as another person coded the same way. Fox News is Fox News. But is it? There is reason to be skeptical. When someone broadly reports that they view Fox News that measure says little about what stimuli – that is, what *causes* – they actually encountered during such viewing. Though two people reporting viewing Fox News report only "nominally identical" experiences that might in practice vary systematically (Rubin 1990, 475). Did they see a particular claim about the Iraq War? Did they see particular on-screen visuals? Did they hear particular arguments about the war? These are components of media that laboratorybased media effects research typical finds to be causally relevant. The self-reported exposure to Fox News does not tell us any of this without assumptions about what that viewership entailed (e.g., in terms of timing and duration) *and* assumptions about what content that time-specific viewership offered (e.g., in terms of information, arguments, issue emphasis, visuals, and so forth). Without these assumptions that experiences are homogeneous (even if viewing Fox News means different things to different people), it is impossible to draw out causal insight from such self-reported exposure. If we want to understand the causal influence of media we have to understand *particularistic* media experiences, exposure or non-exposure to which would generate counterfactual outcomes. Exposure self-reports gloss over this in hopes that a time- or source-based measure proxies for particularistic exposure, yet these measures – despite decades of effort – continue to be coarse, unreliable, and frequently invalid and they regularly become out of date as media landscapes change. Goldman, Mutz, and Dilliplane (2013) argue that despite difficulty of measuring exposure per se and responding to such over-time changes, "measurement consists of the best one can do at any given point in history; we must make do with what is on offer" (651). But there is simply no reason to settle when the entire enterprise is flawed per se.

The Challenges of Measuring Particularistic Media Experiences

Ultimately, what we mean by "media" and "media exposure" and what we can learn about the effects thereof is therefore tied up with the measures we can use to summarize the high-dimensionality of "media" as a concept. Experimental and quasi-experimental approaches to studying media effects avoid these challenges by defining media effects narrowly and studying the effects of isolatable experiences. Media effects in an observational approach, however, typically defines effects more broadly in terms of outcome variation across degrees, amounts, or types of content exposure. This canonical approach poses two insurmountable problems. First, media experiences are infinitely complex and the effort to measure those experiences reduces that complexity in seemingly useful but ultimately futile ways. Our sense of what variation there is in media experience is thus entirely defined by the granularity of our measurement devices. Finer granularity, rather than assuring depth of understanding, simply reveals the further complexity of these more granular slices of media experiences. As new measures identify and attempt to resolve the deficiencies of previous approaches, researchers typically reveal not merely weaknesses in measures—like coarseness—but also inadequate conceptualization—like low-dimensionality—of media experiences in a way that continuously expands the relevant set of complexities that must thereafter be measured. Media effects cannot progress if effects are defined in terms of such complex bundles of causes. Second, though this ever-expanding complexity of conceptualization and measurement might be downplayed in order to obtain partial understandings of media experiences, the complexity and constantly emerging new forms of complexity reveal that media diets involve noncomparable experiences across modes, geographies, time periods, and persons that ultimately limit the extent to which particular simplifications of media complexity that are acceptable in one context can be considered acceptable elsewhere. I term these problems, respectively, the *complexity problem* and the *incommensurability problem* in media measurement and I review them each in turn.

First, to assess media effects, media as a cause must be reducible to a well-defined set of counterfactual experiences that can be measured with minimal bias and maximal precision. But media experiences are not so easily summarized unless experimental methods are deployed. Take, for instance, the hate rally framing example. In that design, media experience is defined narrowly as a single exposure to a single, well-defined message mutually exclusive of exposure to another well-defined message, all else constant. While the broader media landscape is complicated by numerous mediums, numerous sources therein, and an overwhelming abundance of content, "media" becomes narrowly defined so that its effect(s) can be identified. The value of the forward-looking style of causal inference should be immediately obvious because randomization of these alternative experiences and measurement of any relevant outcomes gives immediate and uncomplicated insight into the possible effect(s) of this message. Identifying that effect using observational data is far more challenging. What counts as exposure? Does it have to happen on a particular medium or channel? For how long? How often? How would we know if a given individual saw the message? Would we ask them? What if they've forgotten? It might require a measure of time-specific exposure, plus content analysis of all possible channels or sources through which the information might have been transmitted. What's the universe of such sources? How would we find and categorize them all? Could we track their behavior online using a browser plug-in? What if they opt out? Could we use Nielsen boxes? Yes, if they are already impaneled? Even then, how do we know they were actually attentive? What about radio exposure, or second-hand exposure via social media or interpersonal discussions? Does that count? If it does, how would we know if it occurred? As media landscapes grow more diverse and more numerous, the complexity becomes overwhelming. Like a fractal diagram, the closer we look, the more there is to see. Whatever ruler we hold up to the world reveals that a more precise ruler might grant superior precision and an ultimately, substantively distinct insight. And the answers to these questions about the seemingly infinite complexity of the media landscape only address the challenge of scoring individuals on whether or not they were exposed to a message; we haven't even thought about outcome measurement yet, or holding all else constant.

Researchers have acknowledged this complexity and responded to it by generating measures that respond to previously ignored sources of complexity. As cable and satellite television emerged in the United States, surveys increasingly measured respondents' access to, subscription to, and use of these sources. Similarly as media landscapes diversified from the 1990s onward and as the internet emerged as a key source of information, survey-based measures of media exposure were similarly updated to measure these new sources of landscape complexity. At the same time, the coarseness of self-reported exposure measures also generated innovation in techniques aimed to better capturing complexity, like Nielsen boxes that record television viewing, radio listening devices, media journaling, and passive tracking of web usage. Yet all of the research deploying these more granular measures of media exposure reveal that survey-based measures of media exposure gloss over immense variation in individual media experiences facilitated by the fractionalized and segmented media landscape of the 21st century. That internet users vary in their exposure to political content all the way from zero news stories per campaign to dozens reflects that whatever simplifying value is afforded by coarse, survey-based measures inhibits the use of such measures for meaningful causal inferences. When we attempt to understand the causal effect of broadly defined "media" and measure the putatively causal factor using traditionally coarse measures, we lose insight into precisely the causally interesting within-medium variation. Immense, increasing, and fractal complexity means that "media" defined according to broad notions of "exposure" is an ill-conceived foundation for generating causal inference.

This complexity understandably leads to the kind of oversimplifications that characterize the survey-based, observational media effects literature. Relying on time use measures and coarse summaries of viewership reduces that complexity to continuous measures of time and artificially discrete measures of audience segmentation. As such, it becomes possible to quantitatively process measures of multiple media. Newspaper readership, television viewing, and radio listening, treated as measures of time, can simply be summed. Membership in distinct audiences can be handled with boolean algebra to reduce complexity further into categories like "online news user" or "like-minded news viewer". The audience for Fox in 1998 can be compared to the audience for Fox in 2018. Broadsheet readers in Norway can be compared to broadsheet readers in the United States.

This leads to the second insurmountable challenge. By reducing the complexity of the source experiences, comparability is seemingly simple. But media experiences across sources, mediums, times, and geographies are fundamentally incommensurable. Reading the *New York Times* on September 10th, 2001 is different from reading the *New York Times* on September 12th, 2001. While the editorial line, overall ideology, and issue coverage might typically be stable day-to-day, the content is ultimately different. If we define media effects as differences in potential outcomes in response to well-defined differences in media experiences, then any effect of the *New York Times* today is definitionally distinct from any effect of the *New York Times* tomorrow. The treatment is heterogeneous but we can pretend that it is not because the label for the two treatments is a constant, a result of vagueness rather than comparability. This false equivalence becomes even more obvious when we compare media experiences at the level of medium or the level of time units. An hour of television news may have once meant a relatively homogeneous experience but that is no longer the case in any locale with more than one dominant news source. An hour of internet use might have once conveyed a certain kind of experience but commensurability across people in a specific context is diminished in even a modestly complex media landscape, let alone comparability across time and/or geography. Even comparisons of equivalent time usage of an identical media service, like Netflix or Twitter or Google News, suffers the same problem as interfaces are localized and algorithmically personalized. An hour of this and an hour of that is apples and oranges.

We cannot define let alone speak of the so-called “effect” of such experiences because they are not in fact a singular media experience. The issue is massive variation in the content to which an individual might be exposed if they spend a similar amount of time on even quite similar media outlets.[ix] Ultimately we cannot learn about effects of media unless we can precisely define what we mean by a particular media experience and have measurement tools capable of accurately and precisely measuring whether a given individual had that experience. A possible response is that combinations of survey-based measures of media exposure in tandem with content analysis of media sources might allow for a reduction of complexity by tracing similar experiences across the apparent complexity of sources and exposure patterns. While sidestepping issues of incommensurability by focusing in on a single dimension or feature of media, such approaches multiply rather than reduce apparent complexity by requiring not only precise and unbiased measures of exposure but also precise and unbiased measures of content. It may be that such a mixed-method approach will facilitate observational causal inference because these at least reduce complexity and steer researchers toward definition and measurement of a singular, potentially causal media experience but more work is needed in this area.

Distinguishing Research Goals to Open Multiple Paths Forward

A pessimist might read this whole line of argumentation as a strong case against studying media and media effects at all. The intention, however, is quite different. Rather than abandon the research goals of understanding the role of media in society and politics and rather than abandon observational methods entirely, social scientists must instead come to terms with the reality that multiple goals of research—and thus plural methods for obtaining those distinct goals—are coeval. The goal of obtaining causal inference is something best left to the methods most capable of credibly achieving it.[x] Observational approaches are, by default, inappropriate tools for studying media effects without strong, typically insatiable assumptions. Yet experiments are also limited to providing prospective insight into narrow causal possibilities. Experiments are thus typically deployed inappropriately if in service to any other research goal. And these other research goal are just as, if not more, important. Take for instance the goals of obtain thorough descriptive insights into the abundant content of the ever-expanding media landscape and fine-grained characterizations of the media diets of human populations. These are goals best tackled with methods suited to the purpose. Experimental techniques are not them.

Narrowly defining media effects research as an enterprise of causal inference, as is done here, is meant to highlight that one tool should be primarily deployed in service to that goal. Other closely aligned goals are well-served by alternative approaches. This ideal might be challenging as the social sciences have hit a distinctly confirmatory moment in the history of methodology. The “credibility revolution” has meant that observational methods have almost disappeared from policy evaluation and political economy research in high-profile journals; the fields of political communication and public opinion, which were already heavily experimental, appear to have become even more so. Experiments are no mere fad and the push for trial-like preregistration of analysis plans has pushed the social sciences into even more confirmatory ways of thinking about research methods and the goals of research. Even as qualitative methods have been prominently showcased in recent high-profile work on political behavior and political communication (e.g., Cramer 2017; Nielsen 2012), exploratory, inductive, and thick-descriptive research goals seem to have fallen out of favor in prominent disciplinary outlets. The dominance of the experimental approach has been good for confirmatory research, but it has diverted attention from the multiple, equally valid objectives of social science—that is, *to describe* and *to explain*.

Yet the changing media landscapes that have characterized the period from the 1980s (with the advent of cable television) through the 1990s (with the arrival of the web) through to the 2000s (with the emergence of web 2.0 technologies and social media) are a period more in-demand of thick description than almost any period in the history of social science. A century ago, the Princeton Radio Research Project and later Columbia School of social research aptly understood the societal changes that widely available radio and television would bring and studied the phenomenon in multiple ways including panel surveys (e.g., Lazarsfeld 1940; Berelson et al. 1954; Katz and Lazarsfeld 1960), content analysis (Cantril 1940) and ethnographic work (e.g., Lerner 1958). Just as Lazarsfeld's collaborators were pioneering methods of causal inference about media effects (e.g., Lazarsfeld and Fiske 1938; Lazarsfeld 1940), they were also engaging in deep, descriptive, and exploratory research into media content and media experiences. Now, the set of media have evolved from mere print and broadcast to new digital media that mix textual and audio-visual experiences and are layered by phenomena like dual-screening and the complexly mediated two-step flows of content through social media. Alongside this seismic shift in the diversity of mediums has come a massive escalation in the variety of media alternatives available to the public. The number of alternative media has increased, the number of specific sources has increased, and the sheer volume of content has increased. Description of these changes requires not only cataloging the content of each new outlet using the metrics of more traditional media but also changes in the ways that media are conceptualized and measured. Measures of whether people read a national newspaper or view the evening news intend to capture some metric of political engagement with national politics due to the consistent, measurable, and perhaps predictable content of such outlets. But the question of whether people view YouTube or read news on Facebook communicates almost nothing about what they have experienced. The measures of old media landscapes are inapplicable to new, richer, denser, and more complex media landscapes. Our descriptions grow relatively and ever more thin.

As the complexity of media increases and the rate of change therein accelerates, such thin description leaves ever-greater portions of the political and social world hidden from scholarly attention. Without thick and thorough understanding of this complexity, the hypotheses tested by identification-oriented research will constitute a smaller and smaller portion of the interesting variation in media experiences. What can be done to provide more depth? Both quantitative and qualitative methods are the answer. Computational methods to gather and characterize especially online media content offer the prospect of characterizing media at a scale previously unimaginable, as well as assessing the diversity of media content across times, platforms, geographies, and individuals. Web-tracking methods offer the possibility of studying media exposure with a very high degree of granularity that might be able to meaningfully separate exposure from attention and quantify depth of engagement with particular sources, articles, or issues. Ethnographic, qualitative interviewing, and diary methods offer a quite distinct form of thick description. Just as the production of media has rapidly evolved and changed, necessitating sophisticated approaches to map and characterize the media landscape, the consumption of media necessitates in-depth insight how citizens feel and think about media. A compelling example of this can be found in Toff and Nielsen's (2018) use of in-depth interviews to understand how the seemingly inattentive segments of society understand their own learning processes. Cramer and Toff (2017) use similar methods to demonstrate that what the public considers to be politically consequential knowledge often differs from the type that is measured on political surveys. This form of evidence gathering provides highly useful hypothesis generation. While these kinds of inductive, qualitative methods – like their massive-scale quantitative analogues – will not credibly identify media effects, that is not their ambition. This kind of thick description and exploration is needed more than ever.

Thus the effort here to make a strong case for limiting the study of media effects to experimental approaches must be read as part of a larger advocacy for a more pluralistic social science that is diverse not only in its methods but also in its questions. This idea that different methods suit different goals is familiar, but appeals to mixed methods are typically made with an implicit or explicit goal of triangulation: that is, arriving upon similar answers from different methods (for some discussion, see Seawright 2016). As comparative studies of observational and experimental methods quite consistently show, different methods applied even to identical research questions tend to diverge somewhat in their conclusions (Lalonde 1986; Arceneaux, Gerber, and Green 2006). Triangulation is a flawed way forward. Mixing methods to arrive at different answers to different questions, by contrast, will provide a richness of understanding of media and its effects that no method alone and no method deployed in service to triangulation can provide. Rather than debating whether different methods can triangulate on causal inferences (thus implicitly limiting research attention to confirmatory research), a more fruitful path is discriminating about methods while retaining plurality with respect to research questions.

Conclusion

This chapter has discussed what media effects are and how they might be studied outside of a laboratory-experimental setting. The chapter focused on two major challenges posed by the study of “real-world” media effects—namely, the challenging of operationalizing media and the difficulty of credibly claiming linkages between media and outcomes of interest—and then demonstrated how experimental and quasi-experimental research designs have been effectively used to draw causal inferences about media effects. While the chapter might suggest a degree of fatalism about the media effects literature, all hope is not lost. Indeed, if readers take away one message from this chapter it should be that the question of media effects is too important to be lost in dead ends. We should rely much more on survey and laboratory experiments to understand the mechanisms of media influence, heterogeneity in media effects, and the variety of possible effects media may have. And we should similarly rely much more on field experimental and quasi-experimental designs to understand the direction and magnitude of media influence in real-world settings. Ultimately, we do not need survey-based measures of media exposure to understand media effects, so we should spend much less time, effort, and resources improving them. The opportunity cost is too great and there is too much that we do not yet know.

Yet this is not to diminish observational research in service to closely aligned but qualitatively distinct research objectives. More than ever we need exploratory, observational research for its own sake, to serve the distinct goal of describing media and media experiences and generating theoretical insights that might be tested experimentally. Understanding media through these tools must be a complement to credible methods for obtaining the distinct goal of causal inference. Computational methods, in particular, offer the prospect of systematically characterizing textual media content at a scale previously unimaginable. Digital trace data provided by social media APIs and web-tracking software promises to provide insight into individual-level online experiences that can hardly be understood using aggregated user counts or web traffic statistics or the kinds of survey self-reports that have dominated media research (Lazer and Radford 2017; Menchen-Trevino and Karr 2012; Barbera et al. 2015). At the same time, ethnographic approaches and interview-based methods seem capable of exposing how people think and talk about their experiences of media using language and concepts that cannot necessarily be captured by closed-ended time use questions on survey questionnaires. Discourse analytic and content analysis methods can serve to critique understandings of and interpretations of media content. These objectives are well-served by methods other than experimentation.

But an even more difficult conclusion than this relates to what can be learned from experimental methods. While I have argued that observational methods, by definition, cannot generally provide insight into the causal effects of media and furthermore argued that experimental techniques are uniquely capable in this respect, the reality is that experimental methods are also deficient. They can only provide insight into causal possibilities and only demonstrate or explain phenomena to the extent that they entail experiences, treatments, outcomes, and participants reflective of those of broad interest. They might generalize but it is hard to know how far without extensive, multi-study programs of research. In the end, an individual experiment – regardless of the size and scope of the intervention or the number of participants involved – is never going to be able to comprehensively and generalizably describe the effects of media. But that is an unobtainable ideal that no single instance of any method can obtain. We should learn what we can from experiments—namely, about the possible effects of media—and similarly learn what we can from observational methods—namely about patterns of media content and experience. All of this while acknowledging the fundamental limits to what is knowable and acknowledging that any understanding of media or its effects is prone to be immediately out-of-date.

References

Albertson, Bethany L., and Adria Lawrence. 2009. "After the Credits Roll: The Long-Term Effects of Educational Television on Public Knowledge and Attitudes." *American Politics Research* 37(2): 275–300.

Althaus, Scott L., and David Tewksbury. 2007. Toward a New Generation of Media Use Measures for the ANES Report to the Board of Overseers. Technical report American National Election Studies Board of Governors.

Angrist, Joshua D. 1990. "Lifetime Earnings and the Vietnam Draft Lottery: Evidence from Social Security Administrative Records." *American Economic Review* 80(3): 313–336.

Arceneaux, Kevin, Alan S. Gerber, and Donald P. Green. 2006. "Comparing Experimental and Matching Methods Using a Large-Scale Field Experiment on Voter Mobilization." *Political Analysis* 14(1): 37–62.

Arceneaux, Kevin, and Martin Johnson. 2012. *Changing Minds or Changing Channels? Media Effects in the Era of Viewer Choice*. Chicago, IL: University of Chicago Press.

Barberá, Pablo, John T. Jost, Jonathan Nagler, Joshua A. Tucker, and Richard Bonneau. 2015. "Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber?"*Psychological Science* 26(10): 1531–1542.

Bartels, Larry M. 1993. "Messages Received: The Political Impact of Media Exposure." *American Political Science Review* 87(2): 267–285.

Bennett, W. Lance, and Shanto Iyengar. 2008. "A New Era of Minimal Effects? The Changing Foundations of Political Communication." *Journal of Communication* 58: 707–731.

Berelson, Bernard R., Paul F. Lazarsfeld, and William N. McPhee. 1954. *Voting: A Study of Opinion Formation in a Presidential Campaign*. Chicago, IL: University of Chicago Press.

Cantril, Hadley. 1940. *Gauging Public Opinion*. Princeton University Press.

Chaffee, Steven H., and Joan Schleuder. 1986. "Measurement and Effects of Attention to Media News." *Human Communication Research* 13(1): 76–107.

Clinton, Joshua D., and Ted Enamorado. 2014. "The National News Media's Effect on Congress: How Fox News Affected Elites in Congress." *The Journal of Politics* 76(4): 928—943.

Cramer, Katherine J. 2016. *Politics of Resentment*. University of Chicago Press.

Cramer, Katherine J., and Benjamin J. Toff. 2017. "The Fact of Experience: Rethinking Political Knowledge and Civic Competence." *Perspectives on Politics* 15(3): 754–770.

de Vreese, Claes H., and Hajo G. Boomgaarden. 2006. "Media Message Flows and Interpersonal Communication: The Conditional Nature of Effects on Public Opinion." *Communication Research* 33(1): 19–37.

DellaVigna, Stefano, and Ethan Kaplan. 2007. "The Fox News Effect: Media Bias and Voting." *Quarterly Journal of Economics* 122(3): 1187–1234.

Dilliplane, Susanna. 2011. "All the News You Want to Hear: The Impact of Partisan News Exposure on Political Participation." *Public Opinion Quarterly* 75(2): 287–316.

Dilliplane, Susanna, Seth K. Goldman, and Diana C. Mutz. 2012. "Televised Exposure to Politics: New Measures for a Fragmented Media Environment." *American Journal of Political Science* In press: 32–47.

Druckman, James N., and Thomas J. Leeper. 2012*a*. "Is Public Opinion Stable? Resolving the Micro/Macro Disconnect in Studies of Public Opinion." *Daedalus* 141(4): 50–68.

Druckman, James N., and Thomas J. Leeper. 2012*b*. "Learning More from Political Communication Experiments: Pretreatment and Its Effects." *American Journal of Political Science* 56(4): 875–896.

Druckman, James N., Jordan Fein, and Thomas J. Leeper. 2012. "A Source of Bias in Public Opinion Stability." *American Political Science Review* 106(2): 430–454.

Druckman, James N., Matthew S. Levendusky, and Audrey McLain. 2017. "No Need to Watch: How the Effects of Partisan Media Can Spread via Interpersonal Discussions." *American Journal of Political Science* .

Durante, Ruben, Paolo Pinotti, and Andrea Tesei. 2013. "Voting Alone? The Political and Cultural Consequences of Commercial TV." Unpublished paper, Science Po.

Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya. 2011. "Media and Political Persuasion: Evidence from Russia." *American Economic Review* 101(7): 3253–3285.

Erikson, Robert S., and Laura Stoker. 2011. "Caught in the Draft: The Effects of Vietnam Draft Lottery Status on Political Attitudes." *American Political Science Review* 105(2): 1–17.

Eveland, William P., and Dietram A. Scheufele. 2000. "Connecting News Media Use with Gaps in Knowledge and Participation." *Political Communication* 17(3): 215–237.

Eveland, William P., Andrew F. Hayes, Dhavan V. Shah, and Nojin Kwak. 2005. "Understanding the Relationship Between Communication and Political Knowledge: A Model Comparison Approach Using Panel Data." *Political Communication* 22(4): 423–446.

Eveland, William P., Mihye Seo, and Krisztina Marton. 2002. "Learning From the News in Campaign 2000: An Experimental Comparison of TV News, Newspapers, and Online News." *Media Psychology* 4(4): 353–378.

Eveland, William P., Myiah J. Hutchens, and Fei Shen. 2009. “Exposure, Attention, or ‘Use’ of News? Assessing Aspects of the Reliability and Validity of a Central Concept in Political Communication Research.” *Communication Methods and Measures* 3(4): 223–244.

Feezell, Jessica T. 2017. “Agenda Setting through Social Media: The Importance of Incidental News Exposure and Social Filtering in the Digital Era.” *Political Research Quarterly* , 106591291774489.

Freedman, Paul, and Ken Goldstein. 1999. “Measuring Media Exposure and the Effects of Negative Campaign Ads.” *American Journal of Political Science* 43(4): 1189–1208.

Fridkin, Kim L., Patrick J. Kenney, Sarah Allen Gershon, Karen Shafer, and Gina Serignese Woodall. 2007. “Capturing the Power of a Campaign Event: The 2004 Presidential Debate in Tempe.”*The Journal of Politics* 69(03): 770–785.

Gaines, Brian J., and James H. Kuklinski. 2011. “Experimental Estimation of Heterogeneous Treatment Effects Related to Self-Selection.” *American Journal of Political Science* 55(3): 724–736.

Garrett, R. Kelly, Dustin Carnahan, and Emily K. Lynch. 2013. “A Turn Toward Avoidance? Selective Exposure to Online Political Information, 2004–2008.” *Political Behavior* 35(1): 113–134.

Gelman, Andrew, and Guido W. Imbens. 2013. “Why ask Why? Forward Causal Inference and Reverse Causal Questions.”. NBER Working Paper 19614.

Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. W. W. Norton & Company.

Gerber, Alan S., James G. Gimpel, Donald P. Green, and Daron R. Shaw. 2011. “How Large and Long-Lasting Are the Persuasive Effects of Televised Campaign Ads? Results from a Large Scale Randomized Experiment.” *American Political Science Review* 105(1): 135–150.

Goldman, Seth K., Diana C. Mutz, and Susanna Dilliplane. 2013. “All Virtue Is Relative: A Response to Prior.” *Political Communication* 30(4): 635–653.

Graber, Doris A. 1988. *Processing the News: How People Tame the Information Tide*. Guilford.

Guess, Andrew M. 2014. “Measure for Measure: An Experimental Test of Online Political Media Exposure.” *Political Analysis* 23(1): 59—75.

Hayes, Danny, and Mathieu Turgeon. 2009. “A Matter of Distinction: Candidate Polarization and Information Processing in Election Campaigns.” *American Politics Research* 38(1): 165–192.

Holland, Paul W. 1986. “Statistics and Causal Inference.” *Journal of the American Statistical Association* 81(396): 945–960.

Hovland, Carl I. 1959. “Reconciling Conflicting Results Derived from Experimental and Survey Studies of Attitude Change.” *The American Psychologist* 14(1): 8–17.

Huber, Gregory A., and Kevin Arceneaux. 2007. “Identifying the Persuasive Effects of Presidential Advertising.” *American Journal of Political Science* 51(4): 957–977.

Iyengar, Shanto, and Donald R. Kinder. 1987. *News That Matters: Television and American Opinion*. Chicago, IL: University of Chicago Press.

Jerit, Jennifer, Jason Barabas, William Pollock, Susan Banducci, Daniel Stevens, and Martijn Schoonvelde. 2016. “Manipulated vs. Measured: Using an Experimental Benchmark to Investigate the Performance of Self-Reported Media Exposure.” *Communication Methods and Measures* 10(2-3): 99–114.

Katz, Elihu, and Paul F. Lazarsfeld. 1960 [2005]. *Personal Influence: The Part Played by People in the Flow of Mass Communication*. Brunswick, NJ: Transaction Publishers.

Keele, Luke J., and Rocío Titiunik. 2014. “Geographic Boundaries as Regression Discontinuities.” *Political Analysis* 23(1): 127—155.
King, Gary, Benjamin Schneer, and Ariel White. 2017. “How the news media activate public expression and influence national agendas.” *Science* 358(6364): 776–780.
Krasno, Jonathan S., and Donald P. Green. 2008. “Do Televised Presidential Ads Increase Voter Turnout? Evidence from a Natural Experiment.” *The Journal of Politics* 70(1): 245–261.
Kull, Steven, Clay Ramsay, and Evan Lewis. 2003. “Misperceptions, the Media, and the Iraq War.” *Political Science Quarterly* 118(4): 569–598.
LaLonde, Robert J. 1986. “Evaluating the Econometric Evaluations of Training Programs with Experimental Data.” *American Economic Review* 76(4): 604–620.
Lazarsfeld, Paul F. 1940*a*. “‘Panel’ Studies.” *Public Opinion Quarterly* 4(1): 122–128.
Lazarsfeld, Paul F. 1940*b*. *Radio and the Printed Page*. Ayer Company.
Lazarsfeld, Paul F., and Marjorie Fiske. 1938. “The ‘Panel’ As a New Tool for Measuring Opinion.” *Public Opinion Quarterly* 2(4): 596–612.
Lazer, David, and Jason Radford. 2017. “Data ex Machina: Introduction to Big Data.” *Annual Review of Sociology* 43(1): 19–39.
Leeper, Thomas J. 2017. “How Does Treatment Self-Selection Affect Inferences About Political Communication?” *Journal of Experimental Political Science* 4(1): 21–33.
Lelkes, Yphtach, Gaurav Sood, and Shanto Iyengar. 2015. “The Hostile Audience: The Effect of Access to Broadband Internet on Partisan Affect.” Unpublished paper, University of Amsterdam Unpublished paper.
Lerner, Daniel. 1958. *The Passing of Traditional Society: Modernizing the Middle East*. Free Press of Glencoe.
Martin, Gregory J., and Ali Yurukoglu. 2017. “Bias in Cable News: Persuasion and Polarization.” *American Economic Review* 107(9): 2565–2599.
Menchen-Trevino, Ericka, and Chris Karr. 2012. “Researching Real-World Web Use with Roxy: Collecting Observational Web Data with Informed Consent.” *Journal of Information Technology & Politics* 9(3): 254–268.
Morgan, Stephen L., and Christopher Winship. 2015. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. 2nd ed. New York: Cambridge University Press.
Nelson, Thomas E., Rosalee A. Clawson, and Zoe M. Oxley. 1997. “Media Framing of a Civil Liberties Conflict and Its Effect on Tolerance.” *American Political Science Review* 91(3): 567–583.
Nielsen, Rasmus Kleis. 2012. *Ground Wars: Personalized Communication in Political Campaigns*. Princeton, NJ: Princeton University Press.
Paluck, Elizabeth Levy, Paul Lagunes, Donald P. Green, Lynn Vavreck, Limor Peer, and Robin Gomila. 2015. “Does Product Placement Change Television Viewers’ Social Behavior?” *PLOS ONE* 10(9): e0138610.
Panagopoulos, Costas, and Donald P. Green. 2008. “Field Experiments Testing the Impact of Radio Advertisements on Electoral Competition.” *American Journal of Political Science* 52(1): 156–168.
Potter, W. James. 2011. “Conceptualizing Mass Media Effect.” *Journal of Communication* 61(5): 896–915.
Price, Vincent, and John Zaller. 1993. “Who Gets the News? Alternative Measures of News Reception and Their Implications for Research.” *Public Opinion Quarterly* 57(2): 133–164.
Prior, Markus. 2003. “Any Good News in Soft News? The Impact of Soft News Preference on Political Knowledge.” *Political Communication* 20(2): 149–171.

Prior, Markus. 2007. *Post-Broadcast Democracy: How Media Choice Increases Inequality in Political Involvement and Polarizes Elections*. New York: Cambridge University Press.

Prior, Markus. 2009*a*. "The Immensely Inflated News Audience: Assessing Bias in Self-Reported News Exposure." *Public Opinion Quarterly* 73(1): 130–143.

Prior, Markus. 2009*b*. "Improving Media Effects Research through Better Measurement of News Exposure." *The Journal of Politics* 71(03): 893–908.

Prior, Markus. 2013. "The Challenge of Measuring Media Exposure: Reply to Dilliplane, Goldman, and Mutz." *Political Communication* 30(4): 620–634.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1): 41–55.

Rubin, Donald B. 1978. "Bayesian Inference for Causal Effects: The Role of Randomization." *Annals of Statistics* 6(1): 34–58.

Rubin, Donald B. 1990. "Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies." *Statistical Science* 5(4): 472–480.

Searles, Kathleen, Erika Franklin Fowler, Travis N. Ridout, Patricia Strach, and Katherine Zuber. 2017. "The Effects of Men's and Women's Voices in Political Advertising." *Journal of Political Marketing* , 1–29.

Seawright, Jason. 2016. *Multi-Method Social Science: Combining Qualitative and Quantitative Tools*. Cambridge University Press.

Sekhon, Jasjeet S., and Rocío Titiunik. 2012. "When Natural Experiments Are Neither Natural nor Experiments." *American Political Science Review* 106(1): 1–23.

Sen, Maya, and Omar Wasow. 2016. "Race as a Bundle of Sticks: Designs that Estimate Effects of Seemingly Immutable Characteristics." *Annual Review of Political Science* 19(1): 499–522.

Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2001. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton-Mifflin.

Sinclair, Betsy. 2012. *The Social Citizen: Peer Networks and Political Behavior*. University of Chicago Press.

Slater, Michael D., Catherine E. Goodall, and Andrew F. Hayes. 2009. "Self-Reported News Attention Does Assess Differential Processing of Media Content: An Experiment on Risk Perceptions Utilizing a Random Sample of U.S. Local Crime and Accident News." *Journal of Communication* 59(1): 117–134.

Slothuus, Rune. 2015. "Assessing the Influence of Political Parties on Public Opinion: The Challenge from Pretreatment Effects." *Political Communication* 33(2): In press.

Sovey, Allison J., and Donald P. Green. 2011. "Instrumental Variables Estimation in Political Science: A Readers' Guide." *American Journal of Political Science* 55(1): 188–200.

Tewksbury, David, Scott L. Althaus, and Matthew V. Hibbing. 2011. "Estimating Self-Reported News Exposure Across and Within Typical Days: Should Surveys Use More Refined Measures?" *Communication Methods and Measures* 5(4): 311–328.

Toff, Benjamin, and Rasmus Kleis Nielsen. 2018. "'I Just Google It': Folk Theories of Distributed Discovery." *Journal of Communication* .

Voigtlaender, Nico, and Hans-Joachim Voth. 2014. "Highway to Hitler." Unpublished paper, NBER Working Paper 20150.

[i] This chapter benefited from discussion and feedback from participants at the University of Southern California in November 2017.

[ii] The study of media can also view media as an outcome to be explained—either at the macro level from a supply side perspective or at the micro level in terms of determinants of individual-level demand or exposure—but I set aside these questions for the purposes of this chapter.

[iii] A critique sometimes raised at this point is that even though observational methods risk being subjet to unobserved confounding, they at least provide "externally valid" or "generalizable" insights. This is, however, a canard. Unless an observational method satisfies assumptions that enable causal inference, any supposed "effect" that is identified (e.g., via a regression coefficient) is not generalizable because it is not a valid causal effect estimate to begin with (Morgan and Winship 2015).

[iv] Traditional potential outcomes notation typically omits time subscripts, which can be incorrectly interpreted to mean that within-person, over-time variation in a cause provide direct insight into causal effects. That can be true, but only under a strong assumption that potential outcomes are independent of earlier potential outcomes (in other words, that treatments are non-cumulative). In light of strong evidence of pretreatment dynamics (Druckman and Leeper 2012a; Slothuus 2015), it is not generally appropriate to make that assumption.

[v] We might define x differently, allowing for a variety of other experiences such as non-exposure or exposure to varying mixes of public safety and free speech content, but effects would be defined similarly: the causal effect of any particular kind of coverage would be a difference in individual potential outcomes relative to some specified alternative.

[vi] This is a strict definition of "media effect". Researchers might also be interested in various descriptive such as individuals' interpretations of media, reflections upon media expressed in lay causal language, or subjective perceptions of first- or third-person media effects, but we should not consider those subjective interpretations to be media effect in a strict sense.

[vii] Of course, some media content is actually random. But short of published evaluations of these tests (e.g., Gerber et al. 2011; Panagopoulos and Green 2008; Paluck et al. 2015; King, Schneer, and White 2017), it would be hard to know what is random and what is not.

[viii] There is no reason to believe, without further assumptions, that individuals attentive to different news sources had similar levels of misperceptions in the absence of any news exposure. Self-selection into media and effects of media are empirically indistinguishable in cross-sectional, observational research (see Leeper 2012).

[ix] While researchers have classically distinguished mere exposure from more in-depth attention or reception of content, media effects can occur in the absence of more in-depth engagement and processing of content. What matters then is not what citizens can recall about their experience but the content of those experiences per se. Media effects are defined by stimuli or inputs, not recollection thereof.

[x] Recognizing, of course, that what is considered the most credible method will no doubt change in the future.