Online Field Experiments

Daniel Muise* Jennifer Pan†

March 13, 2018

Abstract

Changes in information communication technology across the Asian region have altered our field substantively and methodologically. The rapid growth of digitized communications allows us to find new purchase in examining questions fundamental to our understanding of communication theories, norms, and practices across Asia. While methods such as text mining and user analytics are increasingly popular among computational scholars, here, we focus on online field experiments, an approach to studying communication that has the potential to overcome many existing obstacles to social scientific inquiry but one that has been used relatively rarely in Asia. In this paper, we discuss what online field experiments are and how they differ from traditional experiments as well as online lab and survey experiments. We show how researchers can go about designing and implementing online field experiments, focusing on issues where online field experiments differ from their traditional counterparts—legal and ethical considerations, construct validity, randomization and spillover, and statistical analyses. Finally we discuss how online field experiments can advance our understanding of communication in Asia by helping researchers to gain insight and make causal inferences on attitudes, behaviors, and interactions that were previously unobservable.

Keywords: Asia, online experiments, field experiments, experimental design

Word Count: 6,116

*Ph.D. Student, Department of Communication, Stanford University, Stanford CA 94305

†Assistant Professor, Department of Communication, Building 120, Room 110 450 Serra Mall, Stanford University, Stanford CA 94305-2050; ł jenpan.com, (650) 725-7326.

1 Introduction

Communication scholars in Asia face research questions and research obstacles unique to the region. Many concepts—media effects, persuasion, censorship—require different theoretical and empirical treatment in geographic context outside of North America and Western Europe. This is especially true for the study of political communication, due to the broad diversity of political systems and historical trajectories across countries, as well as for social norms, which can guide communication behavior (Miike, 2006).

Explicating these conceptual differences is important to the validity of communication research, but comes with operational roadblocks. Macro-level data are often unavailable in Asia because governments or other organizations are unable to collect this information or unwilling to share it publicly. Meso- and micro-level data are often also lacking as infrastructural weaknesses and public reticence can make on-the-ground data collection too difficult, resource-intensive, or time-consuming to be practical. This may be true even if researchers have native-place ties and strong language capabilities.

The surge of communication technology adoption across Asia has the potential to overcome some of these challenges to research. Notably, China has the world's largest online population at 731 million in 2016 (CNNIC, 2017); South Korea arguably has the world's most advanced Internet network (Belson, 2017); Indonesia was dubbed 'the world's most Twitter-addicted nation' very early on (Radwanick, 2010), and Myanmar has recently experienced an unprecedented surge in smartphone adoption (Leong, 2017). All of this connectivity has two important implications for our field: first that communication is expanding rapidly, and second that more interactions are digitally recorded and/or accessible online. Emails, text messages, social media messages, purchasing practices, and search terms are digitally recorded and can sometimes be accessed by researchers through application programming interfaces (APIs) or web-scraping. This tremendous growth in communication data has allowed researchers to make use of computational methods such as text analysis and user analytics to shed light on communication theories and debates (Peng, Zhang, Zhong, & Zhu, 2013; Zhu, Mo, Wang, & Lu, 2011; Zhu, Wang, Qin, & Wu, 2012).

Moreover, the technological changes in Asia are tightly linked to the increasing relevance of the continent, both in the field of communication and on the world stage (So, 2010). International migration, as well as economic gains in Asia, indicate that Asian communication patterns will be increasingly influential around the globe. Studying communication in Asia is not so much a niche field as it is a necessity for understanding society, just as communication research traditions in the U.S. and Europe have long been interpreted.

In this paper, we discuss online field experiments—experiments that leverage platforms or systems that already exist on the Internet to study the motivations and behaviors of individuals, organizations, and even governments. Online field experiments allow us to study phenomena that were previously unobservable and establish causal relationships between quantities of interest. This method is valuable for communication researchers in the Asia Pacific region because it can help shed light on the sub-fields of political communication and public relations, which have been consistently identified as top concerns in the Asian communication literature. Online field experiments can also help move communication research in the Asia Pacific region beyond description to theory testing, and finally, online field experiments offer a method of direct comparison of behaviors that can help contextualize interpretations of communication theory.

This paper proceeds in three sections. Section 2 defines online field experiments, and discusses how they differ from online survey and online lab experiments. In Section 3, we describe how researchers can go about designing and implementing online field experiments. We aid the discussion in this section by drawing from two recent online field experiments conducted in Asia. Finally, we conclude in Section 4 by discussing opportunities that online field experiments provide for communication research in the Asia-Pacific region.

2 Defining Online Field Experiments

The experimental method has a long-standing history in communication research. Experiments allow us to randomly assign a treatment that approximates a variable of interest

to some units (individuals, organizations, offices), while the remaining units receive a control. Randomization helps ensure that unobservable factors, which may affect the outcome, are distributed in similar ways across the treatment and control groups so that we can then estimate outcomes such as the average treatment effect on the treated and the average treatment effect on sub-groups by comparing the average outcomes of the treated and control groups. Experiments contain three key components: a treatment or multiple treatments (x), random assignment of treatment, and measurement of outcomes (y) (Druckman, Green, Kuklinski, & Lupia, 2011).

2.1 Common Types of Experiments

Experiments in social science come in three broad categories: lab experiments, field experiments, and survey experiments (Druckman et al., 2011). In a lab experiment, participants are brought into a laboratory environment where the researcher has near-complete control over the stimuli. Lab experiments are used extensively in social science research, and in communication research (Cacioppo & Petty, 1982; Druckman, 2003; Iyengar, Peters, & Kinder, 1982), including studies focused on the Asia-Pacific region (Henrich et al., 2010; Horiuchi, Komatsu, & Nakaya, 2012).

Survey experiments are based on varied survey design, which are randomly assigned. Examples of survey experiments include vignette experiments, list experiments, endorsement experiments, and randomized response (Blair & Imai, 2012; Blair, Imai, & Lyall, 2014; Bullock, Imai, & Shapiro, 2011; Corstange, 2009; Gaines, Kuklinski, & Quirk, 2006; Glynn, 2010). Communication researchers have used survey experiments to examine racial biases and public opinion (Kuklinski, Cobb, & Gilens, 1997; Sniderman & Piazza, 1995). In Asia, choice experiments have been conducted extensively (Hua, 2009; Jin & Wang, 2006; Ku & Yoo, 2010; Ku, Yoo, & Kwak, 2009; Pek & Jamal, 2011; Sakata, 2007), and more diverse types of survey experiments are emerging as well (Horiuchi, Imai, & Taniguchi, 2007; Naoi & Kume, 2011).

Field experiments are conducted in subjects' natural environment, allowing for a more realistic context but with less control over stimuli (as compared to a lab experiment). In a classic field experiment, Bertrand and Mullainathan (2004) submit resumes to job

openings, varying the race of the applicant on the resume, to examine the effect of racial discrimination on employment. In Chattopadhyay and Duflo (2004), a random subset of village councils elected women as council heads, to measure whether public goods provision differed by leaders' gender. Gerber and Green (2000) randomly assign personal canvassing, telephone calls, and direct mail messages to roughly 30,000 registered voters in New Haven, Connecticut to measure the effect of these strategies on voter turnout. In communication research, Panagopoulos and Green (2008) uses randomly-varied radio advertisements to study electoral competition.[1] In Asia, field experiments have been used to study censorship, government responsiveness, attitudes toward privacy, and political deliberation (J. Chen, Pan, & Xu, 2017; Distelhorst & Hou, 2014, 2017; Hui, Teo, & Lee, 2007; King, Pan, & Roberts, 2014; Kizilcec, Davis, & Cohen, 2017; Kobayashi & Ichifuji, 2015)

2.2 Online Experiments

Any of these three categories of experiments can be conducted over the Internet, and be transformed into an 'online experiment'. For example, participants in a laboratory experiment can engage in the experiment remotely from their own computers, and participants in a survey experiment can be recruited online, taking the survey on their own Internet-connected devices. There are numerous data-collection platforms that allow social scientists to perform online lab and survey experiments. Examples include Amazon's Mechanical Turk (Berinsky, 2004; Berinsky, Quek, & Sances, 2012; Bohannon, 2011; Horton, Rand, & Zeckhauser, 2011; Litman & Abberbock, 2017; Litman, Robinson, & Rosenzweig, 2015; Majima, 2017; Peer, Brandimarte, Samat, & Acquisti, 2017), CrowdFlower, CrowdWorks, ProAcademic, Social-ly, TurkPrime, and Zhubajie (Majima, 2017; Peer et al., 2017; Ruffle & Sosis, 2010).

However, for online survey and lab experiments, the 'online' component of the research essentially refers to participant recruitment and sampling. In contrast, online field experiments tend to capitalize on the web's capabilities and the unique experience of cy-

[1] See Green, Calfano, and Aronow (2014) for description on the state of field experiments in media effects research.

berspace as an environment unto itself. Like online lab and survey experiments, online field experiments take place in online settings, which could include a single existing but proprietary platform, such as Facebook (Aral & Walker, 2014; Bakshy, Eckles, & Bernstein, 2014; Bakshy, Rosenn, Marlow, & Adamic, 2012; Bond et al., 2012; Coppock, Guess, & Ternoviski, 2016; Jones, Bond, Bakshy, Eckles, & Fowler, 2017; Taylor, Bakshy, & Aral, 2013), a custom-designed platform that users independently join for reasons exogenous to the experiment (Centola, 2010, 2011), platforms created in partnership with firms (Hui et al., 2007), or a set of existing and non-proprietary platforms, such as all local government websites in country (J. Chen et al., 2017). Unlike online lab and survey experiments, the treatment and outcome of online field experiments are often native to the online setting. For example, Taylor et al. (2013) studies the effect of online social influences, unique to platforms such as Facebook where content is distributed differentially based on social ties, through an online field experiment. Likewise, King et al. (2014) conduct an online field experiment across 100 social media platforms in China to measure the effect of discussing collective action online on online censorship.

The remainder of this paper focuses on online field experiments, rather than online survey or lab experiments, because online field experiments provide a new approach to studying the effects and affordances of online communication media.[2]

2.3 Online Field Experiments in Asia

Compared to the overall growth of computational methods research, in Asia and elsewhere, only a handful of online field experiments have been conducted in Asia, primarily in China. These experiments include King et al. (2014), which submits social media posts to 100 Chinese social media platforms, randomly varying the content of the posts, to determine what type of content causes online censorship, and also J. Chen et al. (2017), which submits information requests to all county government websites in China, randomly varying the content of the information requests, to determine what factors increase authoritarian responsiveness. Also in China, Kizilcec, Davis, and Cohen (2017) randomly

[2] Also, online lab and survey experiments are relatively similar to traditional lab and survey experiments in their design and implementation.

assigned value relevance affirmation interventions[3] to Chinese learners participating in a Massive Open Online Course (MOOC) to examine the effect of affirmation on student achievement. In Japan, Kobayashi and Ichifuji (2015) randomly assigned Japanese Twitter users to follow a Japanese politician to examine the effect of following on attitudes toward politicians. In Singapore, Hui et al. (2007) partnered with a local firm to randomly assign privacy assurances to survey takers online to examine effects on disclosure preferences. See Appendix 5 for details on each of these studies.

These five existing studies show that online field experiments can shed light on a wide variety of phenomena, including long-standing areas of inquiry in communication (e.g., government responsiveness to societal actors, the effects of political messages on voter behavior), and the affordances and characteristics of communication technologies (e.g., the effect of social media, the motivations of government censorship.)

In the next section, we discuss how researchers can execute online field experiments. Throughout the next section, we will draw from the experiments described in King et al. (2014) and J. Chen et al. (2017). We focus on these two experiments for three main reasons. First, in these two studies, the outcome is measured by actual online behavior, not a survey of user self-reports as is the case in Kobayashi and Ichifuji (2015). Second, both studies include a relatively small number of treatments, which allows for simpler exposition. Finally, we focus on these two studies because they are implemented on non-proprietary platforms, which any researcher can access. Many large online experiments are happening inside large social-networking companies such as Facebook and Twitter (Taylor et al., 2013). The insights from these experiments can have impressive social-scientific implications, but academics generally lack access to these types of experiments due to the proprietary nature of the platforms and data (Kramer, Guillory, & Hancock, 2014).

[3]For details on value relevance affirmation interventions, see Kizilcec, Saltarelli, Reich, and Cohen (2017) .

3 Executing Online Field Experiments

While online field experiments can draw on an extensive methodological literature on experiments, there are considerations unique to online field experiments. In this section, we focus on these points of distinction. We begin by discussing the ethical/legal considerations of conducting online field experiments. We then focus on three areas where researchers are likely to face unique challenges in the design and analysis of online field experiments: construct validity, randomization and spillover, and statistical analysis. Lastly, we address the issue of external validity—a strong concern for many experiments but an area where online field experiments offer an advantage.

3.1 Ethical and Legal Considerations

The first step in any experiment is to ensure that all laws and ethical principles are followed. The Asian context and the online field experimental method merit special attention to this point. Across Asia, local laws and customs can vary broadly. A study that is acceptable in Jakarta might be unacceptable in Aceh. Researchers have an obligation to ensure that participants are not asked or incentivized to engage in locally criminal acts in the course of research tasks. Researchers should have deep in-country expertise or consult with the local experts before proceeding.

Beyond local laws, researchers must consider the ethics of research, and potential costs and benefits for participants. For example, if an experiment in Thailand incidentally collects participants' criticism of the King, that data could be incriminating under lese-majeste laws. Remuneration decisions must be made contextually if incentive payments are used (as in survey and lab experiments). Offering overly-high levels of payment in return for participation can be irrefusable and thus coercive.

Some online field experiments can require deception and waived informed consent, a situation uncommon in traditional lab experiments. For example, J. Chen et al. (2017) used deception and waived informed consent to study responsiveness of local government offices in China to online citizen complaints. It was impossible to obtain consent beforehand because that would have biased the responses of subjects, and participants

were debriefed afterwards to minimize the time commitment of government offices and to minimize any potential negative influence on how government offices dealt with citizen complaints in the future. In general, researchers must take care to minimize sources of risk and harm to participants. We point to McClendon (2012) for a breakdown of harm considerations in using public officials in field experiments, to Panger (2016) for a critical assessment of deception in online experiments, and also to Distelhorst and Hou (2017) for a case-study of ethical considerations within a Chinese communication field experiment.

Although online field experiments are often not as computationally intensive as observational studies, online field experiments may use computational methods in their design or in analysis of results. For example, J. Chen et al. (2017) identified all county government websites in China using automated crawlers, and King et al. (2014) identified 100 Chinese social media platform as the focus of their experiment after measure the volume of posts across thousands of Chinese social media platforms. When researchers are using web-crawlers and custom bots for their work, they should ensure that the software is not harmful, and does not drain bandwidth in the area of interest (Y. Chen & Konstan, 2015). This is especially true in areas with already-slow connections, as in poorer regions.

A final concern specific to online field experiments is the impact of massive-scale interventions. An experiment that can influence the political opinions of millions of people, even in small ways, could sway elections and other political outcomes. For example, in studying online censorship, King et al. (2014) took great care not to fabricate social media content— instead they propagated posts that already existed online in China—in order to minimize influence on the system they set out to study. In the Kizilcec, Davis, and Cohen (2017) study of how value relevance affirmation affects course-completion rates online, the authors took care to increase feelings of affirmation, instead of decreasing feelings of affirmation.

Before embarking on an online field experiment, researchers must understand the legal and ethical implications of their research designs. Researchers should always obtain approval from their university IRBs, and where possible, obtain IRB approval from the institutions of local research partners.

3.2 Construct Validity

Researchers typically adopt experimental approaches when they have a specific theory or hypothesis they want to test. In other words, before designing an experiment, the researcher should have identified a variable (x) that is hypothesized to generate some observable outcome (y). Sometimes this theorized relationship is generated through qualitative research, other times through observational data, and other times from debates in the literature. For example, King et al. (2014) conduct an experiment to determine whether writing about collective action online leads to online censorship in China. It tests a specific theory—that censorship is focused on posts with collective action potential rather than posts that criticize the regime, which was generated through computational methods of text analysis in King, Pan, and Roberts (2013). Construct validity refers to whether the proxy for x actually signals x and not some other feature which could influence outcomes, and whether the measurement of y fully captures, but does not go beyond, the outcome being affected by x (Cronbach & Meehl, 1955).

As with lab and survey experiments, researchers must design a proxy for the variable x that is hypothesized to generate some observable outcome. In contrast to lab and survey experiments, however, researchers conducting online field experiments usually have less freedom, or fewer options, in the design of such proxies. This is because feasibility and ethical constraints often mean that researchers have less control in the field than in a lab. In J. Chen et al. (2017), the researchers hypothesized that three potential factors may motivate autocrats to respond to citizen grievances: the threat of collective action, the threat of tattling to upper-level authorities, and claims of loyalty. J. Chen et al. (2017) proxy these potential sources of authoritarian responsiveness by varying the language of otherwise identical complaints submitted to online government websites. To proxy the threat of collective action, a request for government assistance in obtaining a social welfare benefit is written by the researchers, and at the end of this request, they include the sentence:

> People around me are in a similar situation, they face difficulties, and they also can't get Dibao [the social welfare benefit]. If you can't help, we'll try

> to figure out what we can do together about this situation.

To proxy the threat of tattling to upper level authorities, a request is written that is identical except for replacing the two sentences above with the sentence:

> If this problem cannot be addressed, I'll have to report it to upper-level government officials.

Likewise, to proxy claims of loyalty, a request that is identical except for the sentence:

> I'm a long-standing CCP member, I've always followed the leadership of the Party.

is created.

J. Chen et al. (2017) could have proxied the threat of collective action more explicitly, by mentioning protest or demonstrations, but they did not because such a treatment might have been too politically sensitive for the Chinese context. They could also have proxied claims of loyalty by focusing on government cadres instead of CCP members. Again, they did not because claiming to be a government cadre might have made it more likely that local governments offices would have used government resources and time to investigate the claim. A shortcoming of the threat of collective action treatment in the J. Chen et al. (2017) experiment is that it proxies slightly more than collective action. It mentions collective behavior, ('we'll try to figure out what we can do together'), but it also implies that a relatively larger number of individuals have this problem ('people around me are in a similar situation'). The observed outcomes associated with this treatment could include the effect of threatening collective action and the effect of highlighting a more prevalent social issue. J. Chen et al. (2017) conduct various tests after the experiment to show why the threat of collective action is more likely at work; however, a proxy with greater construct validity would have focused exclusively on the threat of collective action.

Construct validity also encompasses whether the measure of the experimental outcome y adequately proxies the outcome of interest. In J. Chen et al. (2017), the outcome of interest is government responsiveness. J. Chen et al. (2017) measure this outcome by

recording whether the local government office replies to the citizen complaint, the timing of the reply, and the specific content of the reply. These measures of responsiveness represent a conceptualization of responsiveness that is much more limited than classic conceptualizations of responsiveness, which refer to the incorporation of public preferences in policy outcomes (Dahl, 1971). Thus, the experimental outcome of J. Chen et al. (2017) proxies a building block of responsiveness, but does not fully capture the concept of responsiveness.

Achieving construct validity is a particular challenge for online field experiments. Researchers should carefully consider whether their proxy for x and measure of y capture the phenomena they are interested in studying. After initial versions of treatments and outcomes have been developed, they should be thoroughly tested—for example, through interviews and focus groups with respondents similar to those who will receive the treatment in the experiment; through surveys or even survey experiments to measure whether other characteristics are associated with the proxies, and through feedback from other scholars or content area experts. Whenever possible, researchers should seek out opportunities to present and receive feedback on their research designs before conducting their experiment.

3.3 Randomization and Spillover

Another pitfall of online field experiments is the failure of randomization. Randomization could fail in a number of ways. First, it could fail if it is not feasible or if it is not ethical to assign the treatment to some subset of the population and withhold it from others. Second, randomization could fail if those assigned to the treatment group do not receive the treatment (non-compliance). Third, randomization could fail if participants selectively leave the study (attrition). Each of these three pitfalls is common to traditional lab and survey experiments and have been addressed widely elsewhere. In this section, we focus on a failure of randomization distinct to online experiments: *spillover*.

To establish causal inference and to generate an unbiased estimate of the causal quantity of interest—for example the Average Treatment Effect — the stable unit treatment value assumption (SUTVA) must hold. SUTVA includes two components: 1) that the

treatment status of any unit does not affect the potential outcomes of the other units (non-interference), and 2) the treatments for all units are comparable (no variation in treatment). Spillover occurs if the treatment administered to some unit A is received by another unit B for whom the treatment was not intended. Spillover results in the violation of non-interference, and could also lead to variations in treatment. Spillover is a particular concern for online field experiments that take place on social networks where units receiving treatment may be connected.

Spillover between treatment and control groups was a concern for both the King et al. (2014) and J. Chen et al. (2017) studies. King et al. (2014) created accounts on 100 social media sites in China, and submitted posts to these accounts that discussed ongoing collective action events or other trending events happening at the same time. Some posts on these topics were supportive of the Chinese regime, and others were critical. Posts were randomly assigned to accounts, and because several rounds of the experiments were conducted, several posts were made from each account. King et al. (2014) find that posts of collective action are censored, regardless of whether they are critical or supportive of the Chinese regime, but critical posts are not censored (if they are unrelated to collective action). Spillover is a concern if previous censorship of a post made to a particular account influenced subsequent censorship of posts made on that account. In this particular case, spillover can be measured and monitored, and King et al. (2014) find that previous censorship on an account is unlikely to influence the outcomes of the experiment. For J. Chen et al. (2017), spillover would have occurred if neighboring counties realized they were receiving similar complaints, and responded (or did not respond) based not only on the complaint they received but on complaints of neighboring counties. To minimize the risk of spillover, J. Chen et al. (2017) designed their experiment by placing neighboring counties within the same prefecture into different treatment groups because counties in different prefectures are unlikely to communicate.

There are a growing number of method for detecting violations of SUTVA and spillover in networks. Traditionally, these methods have relied on analysis of data after the experiment, e.g., through testing for null hypotheses (Aronow, 2012; Athey, Eckles, & Imbens,

2016; Rosenbaum, 2007), but new research also utilizes alternative experimental designs to test for spillover. For example, Saveski et al. (2017) propose simultaneously running a completely randomized and a cluster-based randomized experiment to identify SUTVA violations due to network effects.

There are also a variety of ways to address the issue of spillover for online field experiments, including 1) using substantive knowledge to avoid spillovers, 2) cluster-based randomization, and 3) statistical methods, such as alternative estimators, to account for spillover. J. Chen et al. (2017) illustrates the first approach, as they randomly assign different treatments to adjacent counties within a Chinese prefecture to avoid spillover effects, leveraging the substantive knowledge that communication between counties across prefectures (and thus spillover) is rare. Experiments using cluster-based randomization are those where units are clustered based on their connections, and treatment conditions are randomly assigned to the cluster-level (Aronow & Middleton, 2013; Eckles, Karrer, & Ugander, 2016; Ugander, Karrer, Backstrom, & Kleinberg, 2013). However, sometimes it is not feasible to avoid spillover by changing the experiment design. For experiments conducted on social networks, e.g., Twitter or Facebook, that are interested in effects on individuals embedded in the social network, researchers can turn to a growing literature focused on causal inference and estimation in settings with spillovers (Aronow, 2012; Athey et al., 2016; Bond et al., 2012; Bowers, Fredrickson, & Panagopoulos, 2012; Christakis & Fowler, 2007; Eckles et al., 2016; Rosenbaum, 2007; Tchetgen & VanderWeele, 2012; Ugander et al., 2013). In many of these studies, spillover effects are modeled with assumptions, usually restrictions, about the network size and structure. For example, Toulis and Kao (2013) rule out effects on friends of friends, and many studies rule out spillover for connections with fewer interactions (Bond et al., 2012; Eckles et al., 2016; Goldenberg, Zheng, Fienberg, & Airoldi, 2010).

Altogether, researchers interested in conducting online field experiments are likely to face unique challenges related to randomization and spillover. However, fortunately, there is growing body of literature on how to detect spillover in online experiments and how to obtain unbiased estimates even in light of network effects.

3.4 Statistical Significance

The above discussion on spillovers, and new methodological approaches in estimation, leads us to a final challenge that scholars conducting online field experiments should be particularly attuned to: statistical significance. Many online field experiments, especially those that take place on proprietary platforms, have extremely large sample sizes (for example see (Bond et al., 2012)). In these large samples, standard t-tests are replaced by their asymptotic form, and critical values—points on the test distribution compared to the test statistic—are drawn from the Normal distribution, which means critical values for testing at the traditional 95% significant level do not increase with sample size.

A number of solutions have been suggested, and here we discuss four. First, researchers should focus on the magnitude of the effect and always report effect sizes. Second, researchers can report other metrics—for example, the Bayesian Information Criterion (BIC) where penalty increases with sample size—along with or instead of p-values. Third, researchers should be transparent about all decisions related to statistical significance—for example, how sample size was determined, whether any data was excluded (if yes, what and why), and all data measured or collected in the study. Finally, a strategy which has gained in popularity is pre-registration of research designs. Before conducting experiments (or other research studies), researchers can create pre-analysis plans that detail what data they will collect, how they will collect data, and how they will analyze data to limit the scope of p-hacking and "fishing" for results. These pre-analysis plans are publicly pre-registered on databases such as EGAP (http://egap.org/) and the Open Science Framework (https://osf.io).

One note is that to date, published online experiments conducted in Asia have relatively small sample sizes because they have not been conducted on large platforms such as Facebook or Twitter, which have relatively low penetration in Asia-Pacific countries such as China. However, even when sample sizes are within standard practices, the strategies discussed above—especially, pre-registration and transparency in reporting—remain beneficial to improving the quality of online field experiments.

3.5 External Validity

Finally, we turn to external validity—the extent to which the causal relationship assessed in the experiment holds over "variations in persons, settings, treatments, and outcomes" (Shadish, Cook, & Campbell, 2002). In other words, if researchers were to conduct the same experiment with other samples, on other platforms, in other online or offline settings, would the results be the same? External validity is a strong concern for many experiments, especially traditional lab experiments, but an area where online field experiments offer an advantage.

By occurring directly in the environment of interest, external validity can be high. Researchers studying online behavior can achieve a high-level of external validity by bringing the aims of the research in line with the experimental design. King et al. (2014) is interested in online censorship in China, and to increase external validity, the authors conduct their experiment on 100 Chinese social media sites, including microblogs, blog sites, and BBS forums. J. Chen et al. (2017) addresses external validity by conducting the experiment on all county-level government websites with complaint forums (over 2,000 counties in total) instead of a random sample or some subset of counties. However, overgeneralization is still a real danger, especially considering the extent of the digital divide in Asia, where the national online population is dissimilar from the national population.

4 Opportunities for Communication Research in Asia

Online field experiments offer tremendous opportunities for communication research in Asia. Compared to regions such as North America or Western Europe, data is much less readily available for researchers working in Asia. Online field experiments take advantage of changes in information communication technology, allowing researchers to access new sources of data and to establish (and measure) causal relationship between variable of interest.

Together, access to new data and measures bring three unique benefits to the study of communication in the Asia Pacific region. First, political communication and public relations have been consistently identified as top concerns in the Asian communication

literature (Kim, Kim, & Choi, 2016; Liu, Liang, & Zheng, 2016; So, 2010). Online field experiments intuitively lend themselves to studies of political opinion and mass-messaging. For example, many of the existing online field experiments conducted in Asia have provided new insights to our understanding of political communication in the region, on topics ranging from censorship to responsiveness (J. Chen et al., 2017; King et al., 2014; Kobayashi & Ichifuji, 2015).

Second, a shortcoming of research on communication in Asia is over-reliance on atheoretical descriptive research (Cheng & Kim, 2010; Lwin & Salmon, 2015; Willnat & Aw, 2004). Online experiments help change this situation since experimentation necessitates a move beyond description, and to tests of theory. Given the political and economic operational roadblocks to performing traditional experiments in Asia and to collecting data rich enough for strong causal inference, online experiments hold particular promise for strengthening theory-driven communication research in the region.

Finally, online experiments offer a method of direct comparison of behaviors and can help overcome some of the challenges of contextualized interpretation of communication theory (Dissanayake, 2009; Goonasekera & Kuo, 2010). For example, a study using Facebook or Twitter as its platform could be replicated almost identically across multiple contexts. While comparison potential may be hampered by nation-specific limitations (e.g., Facebook's limited penetration in China), online experiments can be tailored to contexts in a convincing manner. To that end, J. Chen et al. (2017) can be compared with highly similar responsiveness studies in the United States (Butler, Karpowitz, & Pope, 2012), and other regions of the world (Cleary, 2007; Gilens, 2005; Malesky & Schuler, 2012; Spada & Guimarães, 2013).

Overall, online field experiments are an important approach for improving our understanding of communication in Asia as they help to overcome numerous obstacles associated with traditional experiments and observational methods. Communication systems and communicative behavior are increasingly contained online; with growing internet and mobile penetration, more human experience is capturable over online channels—meaning that experimental approaches can yield increasingly rich and theoretically interesting re-

sults.

5 Appendix: Examples of Online Field Experiments in Asia

Here we describe methodology used in five online field experiments conducted specifically in the Asian context.

1. J. Chen et al. (2017), in "Sources of Authoritarian Responsiveness: A Field Experiment in China", conducted an experiment to measure authoritarian responsiveness among local government officials in China. They made use of thousands of websites set up by local county governments, created to provide a communication channel for citizens in their respective jurisdictions. The ubiquity of these nearly-identical communication channels was due to central government regulations. The authors wrote web-crawling scripts to automatically identify individual websites and sent messages to local officials under the guise of being local citizens, using variations in these messages to act as a treatment. As an outcome, the authors measured whether or not the local official responded, and in what manner. Standard statistical analysis followed data collection, with the unit of analysis being the county.

- Platform: county-level government websites.
- Unit of Analysis: county
- n = 2,103
- Randomization and Treatment: The authors randomly assigned message types to county websites. Adjacent counties within the same prefecture received different treatments, and messages were designed to be relatively mundane. Both strategies were aimed at minimizing the chances of county officials would discuss messages with each other (which could violate the SUTVA assumption). There were three treatment arms, proxying collective action potential, the threat of tattling to higher-level authorities, and party loyalty, as well as a control condition.

- Outcome Measures: There were three outcome variables. Did the message receive a response; how many days did it take to receive a response; content of the response.

- Manipulation Checks: There were no manipulation checks reported in this paper.

2. Hui et al. (2007), in "The Value of Privacy Assurance: an Exploratory Field Experiment", partnered with a Singaporean market-research firm to examine determinants of privacy preferences online. The local firm recruited participants via mass email under the premise of a standard market research survey. Interested participants were directed to the firm's actual website. Before taking the survey, participants were randomly assigned into one of three levels of privacy assurance, one of nine levels of incentive payment for completing the survey, and one of twenty levels of sensitive information requested in the survey. Participants were away of the privacy assurances, monetary compensation, and informational requests at the outset. The outcome was the amount of private information that participants chose to disclose in the survey.

- Platform: The website of a market-research firm (with email recruitment from the firm's pre-existing pool).

- Unit of Analysis: individuals

- n = 109

- Randomization and Treatment: The factorial treatment condition combinations were generated according to a uniform distribution, and these treatment conditions were then randomly assigned to participants as they arrived at the site. The treatment conditions were three levels of privacy assurance, one of nine levels of incentive payment for completing the survey, and one of twenty levels of sensitive information requested in the survey.

- Outcome Measures: The outcome measure, private information disclosure, was intended to capture participants' privacy tolerance. To check this assumption, the authors included a question that asked participants about their general truthfulness

in disclosing information; low truthfulness would imply that the information provided may be less sensitive than that of a wholly truthful individual. Controlling for self-reported truthfulness did not change the results.

- Manipulation Checks: To check that the sensitivity treatments were administered as expected, the authors asked participants to rate each question on how sensitive they felt it was. Also, the authors checked that participants correctly interpreted the privacy assurances in a post-survey.

3. King et al. (2014), in "Reverse-engineering Censorship in China: Randomized Experimentation and Participant Observation", study determinants of government-backed censorship behavior in Chinese social media. They generated thousands of social media posts in Chinese and posted them online. Posts varied in whether or not they discussed events with collective action potential and whether they were supportive or critical of the government. They then used computational methods to determine whether or not the post was removed.

- Platform: 100 top social media sites in China
- Unit of Analysis: Social media posts
- n = 1200
- Randomization and Treatment: Posts were randomly assigned to social media websites.
- Outcome Measures: The outcome was whether or not a post was censored. The censorship status of each post was determined computationally.
- Manipulation Checks: There were no manipulation checks reported in this paper.

4. Kizilcec, Davis, and Cohen (2017), in "Towards Equal Opportunities in MOOCs: Affirmation Reduces Gender & Social-Class Achievement Gaps in China", introduced a randomly-assigned value relevance affirmation intervention (Kizilcec, Saltarelli, et al., 2017) to Chinese students in an online English-learning course. The object of interest

was the impact of Social Identity Threat (SIT) on engagement and attrition rates, with affirmation introduced to reduce SIT. The authors collaborated with an instructor of the online course to facilitate the study.

- Platform: A MOOC (Massive Open Online Course) operated by Tsinghua University, on the MOOC-operator platform XuetangX.
- Unit of Analysis: Individuals (students)
- n = 1990
- Randomization and Treatment: Within the class, students were randomly assigned to complete a survey module embedded with either a value relevance affirmation exercise or tips on studying. The authors did not explicitly address spillover concerns, though it's reasonable to assume that MOOC students were not exposed to others students' stimuli.
- Outcome Measures: The outcome variables were course completion, grade, and number of attempted assignments, all used as proxies for educational outcomes in online learning environments.
- Manipulation Checks: There were no manipulation checks reported in this paper; in this case, as with other studies of this type, manipulation checks may have triggered reactance, i.e., negative feelings among students.

5. Kobayashi and Ichifuji (2015), in "Tweets that Matter: Evidence from a Randomized Field Experiment in Japan", study the impact of politicians' social media campaigns on voter preferences. They randomly assigned voters to follow politicians on Twitter; control-group participants followed two placebo politicians, while treatment-group participants followed an additional politician, Toru Hashimoto. The outcome measure was whether there was change in feeling toward Hashimoto following the study period. The authors ensured compliance (exposure to politicians' exogenous Tweets) by continuously checking the Twitter API.

- Platform: Twitter.

- Unit of Analysis: individual Twitter users
- n = 773
- Randomization and Treatment: Individuals were randomly assigned to either be exposed or not exposed to tweets from Toru Hashimoto. The authors did not explicitly control for spillover effects, but were able to ensure that participants stayed in their assigned groups.
- Outcome Measures: Pre- and post-surveys were used to measure changes in opinion toward Toru Hashimoto. While survey measures are efficient and ubiquitous, the validity of self-reports faces a long history of criticism in the social sciences. In this study, administering surveys was a reasonable and appropriate method, but we suggest that online field experiments capitalize on the chance to measure behaviors directly. For example, the authors of this study might have measured voter turnout, changes in the political content consumed on Twitter, or increased engagement in political discussions online.
- Manipulation Checks: The authors used the Twitter API to determine if tweets had been seen by participants.

References

Aral, S., & Walker, D. (2014). Tie strength, embeddedness, and social influence: A large-scale networked experiment. *Management Science*, *60*(6), 1352–1370.

Aronow, P. (2012). A general method for detecting interference between units in randomized experiments. *Sociological Methods & Research*, *41*(1), 3–16.

Aronow, P., & Middleton, J. (2013). A class of unbiased estimators of the average treatment effect in randomized experiments. *Journal of Causal Inference*, *1*(1), 135–154.

Athey, S., Eckles, D., & Imbens, G. (2016). Exact p-values for network interference. *Journal of the American Statistical Association*(just-accepted).

Bakshy, E., Eckles, D., & Bernstein, M. S. (2014). Designing and deploying online field experiments. In *Proceedings of the 23rd international conference on world wide web* (pp. 283–292).

Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012). The role of social networks in information diffusion. In *Proceedings of the 21st international conference on world wide web* (pp. 519–528).

Belson, D. (2017). State of the internet report, q1 2017. *ACM SIGOPS Operating Systems Review*, *10*(1).

Berinsky, A. (2004). *Silent voices: Public opinion and political participation in america*. Princeton, NJ: Princeton University Press.

Berinsky, A., Quek, K., & Sances, M. (2012). Conducting online experiments on mechanical turk. *Newsletter of the APSA Experimental Section*, *3*(1), 2–9.

Bertrand, M., & Mullainathan, S. (2004). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *The American Economic Review*, *94*(4), 991–1013.

Blair, G., & Imai, K. (2012). Statistical analysis of list experiments. *Political Analysis*, *22*, 47-77.

Blair, G., Imai, K., & Lyall, J. (2014). Comparing and combining list and endorsement experiments: Evidence from afghanistan. *American Journal of Political Science*, *58*(4), 1043–1063.

Bohannon, J. (2011). Social science for pennies. *Science*, *334*(6054), 307–307.

Bond, R., Fariss, C., Jones, J., Kramer, A., Marlow, C., Settle, J., & Fowler, J. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, *489*(7415), 295–298.

Bowers, J., Fredrickson, M., & Panagopoulos, C. (2012). Reasoning about interference between units: A general framework. *Political Analysis*, *21*(1), 97–124.

Bullock, W., Imai, K., & Shapiro, J. (2011). Statistical analysis of endorsement experiments: Measuring support for militant groups in pakistan. *Political Analysis*, *19*(4), 363-384.

Butler, D. M., Karpowitz, C. F., & Pope, J. C. (2012). A field experiment on legislators home styles: Service versus policy. *The Journal of Politics*, *74*(2), 474–486.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of personality and social psychology*, *42*(1), 116.

Centola, D. (2010). The spread of behavior in an online social network experiment. *science*, *329*(5996), 1194–1197.

Centola, D. (2011). An experimental study of homophily in the adoption of health behavior. *Science*, *334*(6060), 1269–1272.

Chattopadhyay, R., & Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in india. *Econometrica*, *72*(5), 1409–1443.

Chen, J., Pan, J., & Xu, Y. (2017). Sources of authoritarian responsiveness: A field experiment in china. *American Journal of Political Science*, *60*(2), 383-400.

Chen, Y., & Konstan, J. (2015). Online field experiments: a selective survey of methods. *Journal of the Economic Science Association*, *1*(1), 29–42.

Cheng, H., & Kim, K. K. (2010). Research on advertising in asia: a critical analysis of the articles published in major advertising and communication journals, 1990–2009. *Asian Journal of Communication*, *20*(2), 248–263.

Christakis, N., & Fowler, J. (2007). The spread of obesity in a large social network over 32 years. *n engl j med*, *2007*(357), 370–379.

Cleary, M. (2007, April). Electoral competition, participation, and government responsiveness in mexico. *American Journal of Political Science*, *51*(2), 283299.

CNNIC. (2017). *The 39th china internet development statistics report.* http://www.cnnic.net.cn/.

Coppock, A., Guess, A., & Ternoviski, J. (2016). When treatments are tweets: A network mobilization experiment over twitter. *Political Behavior*, *38*(1), 105-128.

Corstange, D. (2009). Sensitive questions, truthful answers? modeling the list experiment with listit. *Political Analysis*, *17*(1), 4563.

Cronbach, L., & Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281.

Dahl, R. (1971). *Polyarchy: Participation and opposition*. Yale University Press.

Dissanayake, W. (2009). The production of asian theories of communication: Contexts and challenges. *Asian Journal of Communication*, *19*(4), 453–468.

Distelhorst, G., & Hou, Y. (2014). Ingroup bias in official behavior: A national field experiment in china. *Quarterly Journal of Political Science*, *9*(2), 203–230.

Distelhorst, G., & Hou, Y. (2017). Constituency service under nondemocratic rule: Evidence from china. *Journal of Politics*, *79*(3).

Druckman, J. N. (2003). The power of television images: The first kennedy-nixon debate revisited. *The Journal of Politics*, *65*(2), 559–571.

Druckman, J. N., Green, D. P., Kuklinski, J. H., & Lupia, A. (2011). An introduction to core concepts. *The Cambridge Handbook of Experimental Political Science*.

Eckles, D., Karrer, B., & Ugander, J. (2016). Design and analysis of experiments in networks: Reducing bias from interference. *Political Behavior*, *38*(1), 105-128.

Gaines, B. J., Kuklinski, J. H., & Quirk, P. J. (2006). The logic of the survey experiment reexamined. *Political Analysis*, *15*(1), 1–20.

Gerber, A. S., & Green, D. P. (2000). The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment. *American political science review*, *94*(3), 653–663.

Gilens, M. (2005). Inequality and democratic responsiveness. *Public Opinion Quarterly*, *69*(5), 778-796.

Glynn, A. (2010). *What can we learn with statistical truth serum? design and analysis of the list experiment* (Tech. Rep.). Harvard University. Retrieved from http://scholar.harvard.edu/aglynn/files/sts.pdf

Goldenberg, A., Zheng, A., Fienberg, S., & Airoldi, E. (2010). A survey of statistical network models. *Foundations and Trends in Machine Learning*, *2*(2), 129–233.

Goonasekera, A., & Kuo, E. (2010). Foreward. *Asian Journal of Communication*, *20*(2), 230–247.

Green, D. P., Calfano, B. R., & Aronow, P. (2014). Field experimental designs for the study of media effects. *Political Communication*, *31*(1), 168–180.

Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., ... others (2010). Markets, religion, community size, and the evolution of fairness and punishment. *science*, *327*(5972), 1480–1484.

Horiuchi, Y., Imai, K., & Taniguchi, N. (2007). Designing and analyzing randomized experiments: Application to a japanese election survey experiment. *American Journal of Political Science*, *51*(3), 669–687.

Horiuchi, Y., Komatsu, T., & Nakaya, F. (2012). Should candidates smile to win elections? an application of automated face recognition technology. *Political Psychology*, *33*(6), 925–933.

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, *14*(3), 399–425.

Hua, G. (2009). An experimental investigation of online banking adoption in china. *The Journal of Internet Banking and Commerce*, *14*(1), 1–12.

Hui, K.-L., Teo, H. H., & Lee, S.-Y. T. (2007). The value of privacy assurance: an exploratory field experiment. *Managment Information Systems Quarterly*, 19–33.

Iyengar, S., Peters, M. D., & Kinder, D. R. (1982). Experimental demonstrations of the not-so-minimal consequences of television news programs. *American Political Science Review*, *76*(4), 848–858.

Jin, J., & Wang, Z. (2006). Choice experiment method and its application to solid waste management in macao. *Huan jing ke xue= Huanjing kexue/[bian ji, Zhongguo ke xue yuan huan jing ke xue wei yuan hui" Huan jing ke xue" bian ji wei yuan hui.]*, *27*(4), 820–824.

Jones, J. J., Bond, R. M., Bakshy, E., Eckles, D., & Fowler, J. H. (2017). Social influence and political mobilization: Further evidence from a randomized experiment in the 2012 us presidential election. *PloS one*, *12*(4), e0173851.

Kim, D., Kim, S.-Y., & Choi, M.-I. (2016). The pivotal role of ajc in the growth of communication research in asia: A semantic network analysis. *Asian Journal of Communication*, *26*(6), 626–645.

King, G., Pan, J., & Roberts, M. (2013, May). How censorship in china allows government criticism but silences collective expression. *American Political Science Review*, *107*(2), 1-18.

King, G., Pan, J., & Roberts, M. (2014). Reverse-engineering censorship in china: Randomized experimentation and participant observation. *Science*, *345*(6199), 1-10.

Kizilcec, R. F., Davis, G. M., & Cohen, G. L. (2017). Towards equal opportunities in moocs: Affirmation reduces gender & social-class achievement gaps in china. In *Proceedings of the fourth (2017) acm conference on learning at scale* (pp. 121–130).

Kizilcec, R. F., Saltarelli, A. J., Reich, J., & Cohen, G. L. (2017). Closing global achievement gaps in moocs. *Science*, *355*(6322), 251–252.

Kobayashi, T., & Ichifuji, Y. (2015). Tweets that matter: Evidence from a randomized

field experiment in japan. *Political Communication*, *32*(4), 574–593.

Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, *111*(24), 8788–8790.

Ku, S.-J., & Yoo, S.-H. (2010). Willingness to pay for renewable energy investment in korea: A choice experiment study. *Renewable and Sustainable Energy Reviews*, *14*(8), 2196–2201.

Ku, S.-J., Yoo, S.-H., & Kwak, S.-J. (2009). Willingness to pay for improving the residential waste disposal system in korea: A choice experiment study. *Environmental management*, *44*(2), 278–287.

Kuklinski, J., Cobb, M., & Gilens, M. (1997). Racial attitudes and the "new south". *The Journal of Politics*, *59*(2), 323-349.

Leong, L. (2017). Mobile myanmar: The development of a mobile app culture in yangon. *Mobile Media & Communication*, *5*(2), 139–160.

Litman, L., & Abberbock, T. (2017). Turkprime. com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior research methods*, *49*(2), 433–442.

Litman, L., Robinson, J., & Rosenzweig, C. (2015). The relationship between motivation, monetary compensation, and data quality among us-and india-based workers on mechanical turk. *Behavior research methods*, *47*(2), 519–528.

Liu, X., Liang, X., & Zheng, P. (2016). A content analysis of research on china in top-ranked communication journals from 1995 to 2014: a comparison with other asian countries. *Asian Journal of Communication*, *26*(6), 605–625.

Lwin, M. O., & Salmon, C. T. (2015). A retrospective overview of health communication studies in asia from 2000 to 2013. *Asian Journal of Communication*, *25*(1), 1–13.

Majima, Y. (2017). The feasibility of a japanese crowdsourcing service for experimental research in psychology. *SAGE Open*, *7*(1), 2158244017698731.

Malesky, E., & Schuler, P. (2012). Nodding or needling: Analyzing delegate responsiveness in an authoritarian parliament. *American Political Science Review*, *104*(3), 482-502.

McClendon, G. H. (2012). Ethics of using public officials as field experiment subjects. *Newsletter of the APSA Experimental Section*, *3*(1), 13–20.

Miike, Y. (2006). Non-western theory in western research? an asiacentric agenda for asian communication studies. *The Review of Communication*, *6*(1-2), 4–31.

Naoi, M., & Kume, I. (2011). Explaining mass support for agricultural protectionism: Evidence from a survey experiment during the global recession. *International Organization*, *65*(4), 771–795.

Panagopoulos, C., & Green, D. P. (2008). Field experiments testing the impact of radio advertisements on electoral competition. *American Journal of Political Science*, *52*(1), 156–168.

Panger, G. (2016). Reassessing the facebook experiment: critical thinking about the validity of big data research. *Information, Communication & Society*, *19*(8), 1108–1126.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163.

Pek, C.-K., & Jamal, O. (2011). A choice experiment analysis for solid waste disposal option: A case study in malaysia. *Journal of environmental management*, *92*(11), 2993–3001.

Peng, T.-Q., Zhang, L., Zhong, Z.-J., & Zhu, J. J. (2013). Mapping the landscape of internet studies: Text mining of social science journal articles 2000–2009. *New Media & Society*, *15*(5), 644–664.

Radwanick, S. (2010). Indonesia, brazil and venezuela lead global surge in twitter usage. *ComScore.com*.

Rosenbaum, P. R. (2007). Interference between units in randomized experiments. *Journal of the American Statistical Association*, *102*(477), 191–200.

Ruffle, B. J., & Sosis, R. H. (2010). *Do religious contexts elicit more trust and altruism? an experiment on facebook.*

Sakata, Y. (2007). A choice experiment of the residential preference of waste management services–the example of kagoshima city, japan. *Waste Management*, *27*(5), 639–644.

Saveski, M., Pouget-Abadie, J., Saint-Jacques, G., Duan, W., Ghosh, S., Xu, Y., & Airoldi, E. (2017). Detecting network effects: Randomizing over randomized experiments. In *Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining* (pp. 1027–1035).

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Wadsworth Cengage learning.

Sniderman, P. M., & Piazza, T. L. (1995). *The scar of race*. Harvard University Press.

So, C. Y. (2010). The rise of asian communication research: A citation study of ssci journals. *Asian Journal of Communication*, *20*(2), 230–247.

Spada, P., & Guimarães, F. d. S. (2013). Investigating elite behavior through field experiment in brazil: do candidates answer more to core or swing voters? *Brazilian Political Science Review*, *7*(1), 56–73.

Taylor, S. J., Bakshy, E., & Aral, S. (2013). Selection effects in online sharing: Consequences for peer adoption. In *Proceedings of the fourteenth acm conference on electronic commerce* (pp. 821–836).

Tchetgen, E., & VanderWeele, T. (2012). On causal inference in the presence of interference. *Statistical methods in medical research*, *21*(1), 55–75.

Toulis, P., & Kao, E. (2013). Estimation of causal peer influence effects. In *International conference on machine learning* (pp. 1489–1497).

Ugander, J., Karrer, B., Backstrom, L., & Kleinberg, J. (2013). Graph cluster randomization: Network exposure to multiple universes. In *Proceedings of the 19th acm sigkdd international conference on knowledge discovery and data mining* (pp. 329–337).

Willnat, L., & Aw, A. J. (2004). Political communication in asia: Challenges and opportunities. *Handbook of political communication research*, 479–503.

Zhu, J., Mo, Q., Wang, F., & Lu, H. (2011). A random digit search (rds) method for sampling of blogs and other user-generated content. *Social Science Computer Review*, *29*(3), 327–339.

Zhu, J., Wang, X., Qin, J., & Wu, L. (2012). Assessing public opinion trends based on user search queries: validity, reliability, and practicality. In *The annual conf. of the*

world association for public opinion research (pp. 1–7).