How Does Treatment Self-Selection Affect Inferences About Political Communication?

Thomas J. Leeper*
Department of Political Science
Aarhus University
thosjleeper@gmail.com

June 10, 2015

Ecological validity is vital to the importance of experimental findings because experiments that provide too stylized, too contrived, or too artificial an experience for participants may not speak to any real world phenomenon of interest. One recent concern related to ecological validity is the idea of treatment self-selection: if individuals in the real world self-select their own treatments, to what extent does a randomized experiment constitute a realistic exploration of that phenomenon? To explore this question, the present research tests how individuals respond to arguments about a political policy that are either randomly assigned or self-selected by the participants. The findings suggest that treatment randomization masks effect heterogeneity between individuals that would have selected different messages if given the option. Yet such treatment selections are themselves complicated, revealing further challenges for realistically studying treatments prone to self-selection. The findings have important implications for experimental research on human behavior and for research on selective exposure, motivated reasoning, and political communication.

*An earlier version of this paper was presented at the Midwest Political Science Association 2012 Annual Meeting, Chicago, IL and the International Society for Political Psychology 2012 Annual Meeting, Chicago, IL. Thanks are due to John Brehm, Jamie Druckman, Samara Klar, Jeff Mondak, and Marco Steenbergen for feedback on various stages of the project, and to Jamie, Fay Lomax Cook, and Toby Bolsen for sharing survey data from the $t1$ panel wave.

Experiments are seen as the gold standard of social science research because the randomized assignment to a treatment provides unparalleled internal validity, ensuring that differences in outcomes between groups are due to the treatment alone. While experiments are sometimes criticized for lacking external validity (representativeness with respect to units, treatments, outcomes, or settings; see McDermott 2011; Shadish, Cook, and Campbell 2001), another important but less studied concern is the notion of ecological validity: i.e., the "realism" of the experiment. Ecological validity is vital to experimental design because studies that involve too stylized or too contrived an experience for participants may not speak to any real world political phenomenon of interest.

One recent concern about ecological validity relates to the notion of treatment self-selection (Lau and Redlawsk 2006; Gaines and Kuklinski 2011*b*,*a*; Druckman, Fein, and Leeper 2012; Arceneaux and Johnson 2012). If in the real world, individuals self-select their own "treatments," to what extent does the randomized exposure within an experiment constitute a realistic exploration of that phenomenon? This is particularly important in the study of political communication, where real-world exposure to messages is often (and perhaps increasingly) self-selected. To explore this question, the present research uses a population-based survey experiment to test how individuals respond to arguments about a political policy question that are either randomly assigned or self-selected by the participants. The findings suggest that treatment randomization masks effect heterogeneity across individuals inclined to select alternative messages. Furthermore, a pretreatment manipulation of attitude importance in the experiment modified participants' choice of treatments and led to variations in apparent treatment effects, suggesting that the inferences from choice-based experiments hinge on which respondents self-select into which treatment. The results speak to political psychological research on selective exposure, motivated reasoning, and political communication, and more generally to the use of experiments for studying processes defined by real-world treatment self-selection.

Studying the Effects of Self-Selected Treatments

Arguments conveyed by media and political elites are widely seen as an important source of political information that citizens can use when learning about, evaluating, and forming preferences about politics (Disch 2011; Chong and Druckman 2007). Much of the evidence for the influence of elite communications from randomized experiments that expose participants to different messages and then measure information effects on outcomes such argument evaluations, opinions, issue importance, and information-seeking (see, for example, Ansolabehere et al. 1994; Arceneaux and Johnson 2012; Berinsky and Kinder 2006; Brewer 2003; Brewer and Gross 2005; Iyengar and Kinder 1987; Levendusky 2013; Miller and Krosnick 2000; Nelson, Clawson, and Oxley 1997; Petty and Cacioppo 1986; Tichenor, Donohue, and Olien 1970). Questions have been raised, however, about the extent to which such experiments provide an adequate empirical test of real-world information exposure given that, in Hovland's words: "In an experiment the audience on whom the effects are being evaluated is one which is fully exposed to the communication. On the other hand, in naturalistic situations with which surveys are typically concerned, the outstanding phenomenon is the limitation of the audience to those who expose themselves to the communication" (Hovland 1959, 9; Bennett and Iyengar 2008, 724). Indeed, despite early skepticism (Sears and Freedman 1967), it is now a relatively uncontroversial claim that citizens selectively expose themselves to political information (Bennett and Iyengar 2008; Bolsen and Leeper 2013; Stroud 2011; Kim 2007, 2009; Iyengar and Hahn 2009; Iyengar et al. 2008; Garrett 2009*b*,*a*; Garrett, Carnahan, and Lynch 2013; Feldman et al. 2011; Smith, Fabrigar, and Norris 2008) and a recent meta-analysis suggests that *political* selective exposure is particularly prevalent (Hart et al. 2009).

An open question therefore relates to the extent to which the randomized experiment is an ecologically valid representation of real-world information exposure and whether or not randomized experiments provide causal inferences that are useful for understanding real-world

processes characterized by treatment self-selection. To understand this requires unifying randomized exposure and message self-selection within a single experimental investigation and comparing the insights from each portion of the design. Gaines and Kuklinski (2011*a*) label such an empirical merger a "hybrid experiment," where one-half of participants participate in a randomized experiment and the other half participate in an observational study involving treatment self-selection. They argue that the hybrid design provides particularly useful insights into treatment effect heterogeneity (i.e., the effects for treatment self-selectors and treatment avoiders). It is also useful for studying the extent to which a randomized experiment provides comparable insights to an observational study (see also Steiner et al. 2010). The present research builds on this work by employing a hybrid design with three additional features: (1) a pretreatment manipulation of attitude importance that is meant to modify the degree to which respondents choose treatment messages consistent with their prior attitudes, (2) a two-wave panel design that allows for a clean measure of prior opinions, which are the theorized cause of message selection, and (3) a nationally representative sample of adult participants. With this design it is possible to examine how individuals in the study choose the treatment they receive and, in turn, how the apparent effects of a message compare under alternative conditions in which it is self-selected (as in the real world) versus randomly assigned (as in typical experimental studies).

Theoretically, I build from research on motivated reasoning (Kunda 1990; Lodge and Taber 2000; Nickerson 1998; Taber and Lodge 2006) to argue that individuals are likely to engage in attitude-congruent selective exposure (i.e., choose messages congruent with their prior opinions) and that this bias in information selection is likely to be most pronounced among those with high attitude importance (Holbrook et al. 2005; Taber and Lodge 2006). Similarly, while individuals are likely to see attitude-congruent arguments as stronger and more effective than incongruent arguments (Ditto et al. 1998), this effect should be more pronounced among those with high attitude importance, due to their stronger desire to defend their prior. Individuals high in importance should also be less likely to change their

attitudes in the face of incongruent information than those low in attitude importance. High attitude importance should also increase information-seeking in general, as has been demonstrated in prior work (Krosnick 1990; Holbrook et al. 2005).

Experimental Design

I test for this comparison of self-selected and randomized exposure in a nationally representative, online, two-wave survey experiment regarding opinions on the issue of "renewable energy portfolio" standards, which require electrical utilities to produce energy from renewable resources.[1] A total of 879 respondents completed both waves and analysis is restricted to these respondents.[2] The first wave of data was collected in Summer 2010 (hereafter, $t1$) and involved collection of demographic covariates and baseline attitudes. The second, experimental wave was collected in Spring 2011 (hereafter, $t2$). The two-wave design is advantageous because it provides a clean measure of $t1$ opinion, avoiding accessibility or consistency biases into respondents' behavior during the survey experiment.[3]

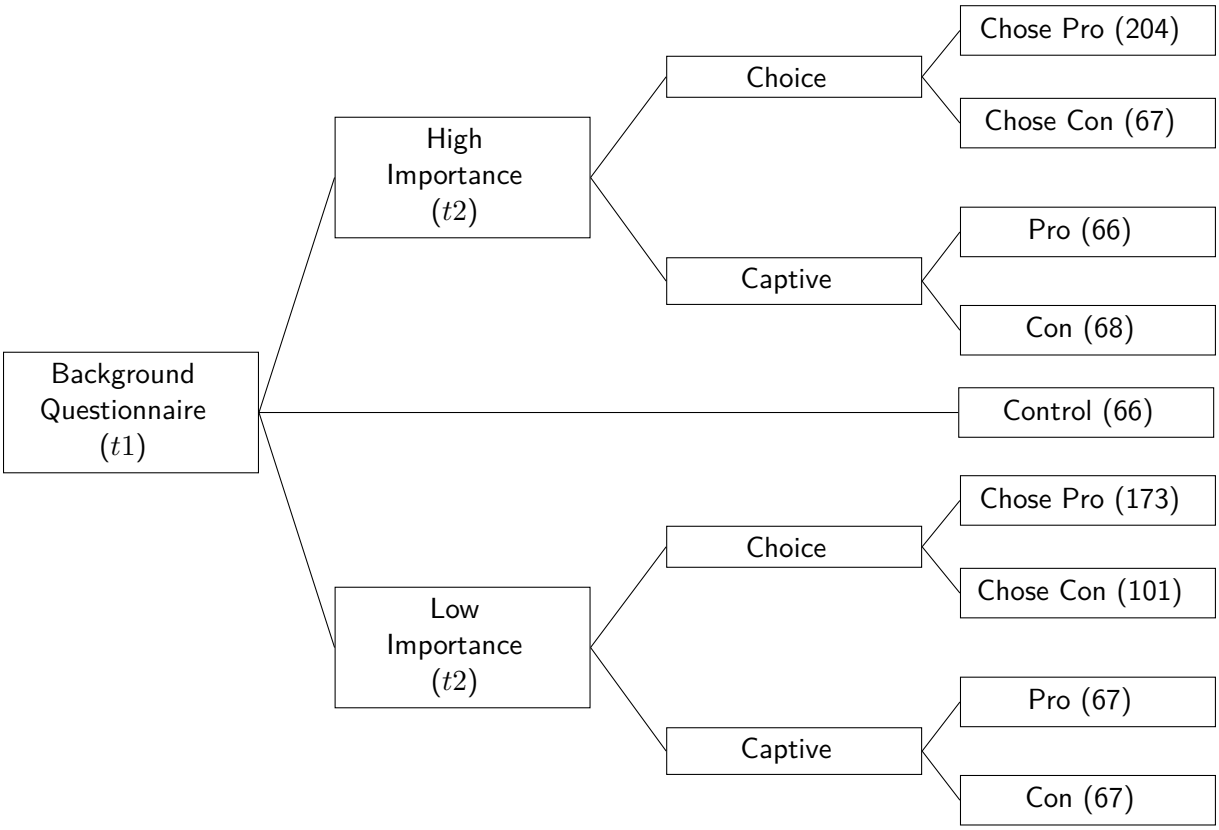
The $t2$ experiment involved three primary manipulations: a manipulation of attitude importance, the direction of the energy policy argument, and whether that argument was self-selected by respondents or captively (i.e., randomly) received. Figure 1 provides a visual summary of the design, along with treatment group sample sizes. The first manipulation involves modifying the apparent personal impact of the energy proposal. This serves as

[1]U.S. Department of Energy, 2003, "Analysis of a 10-percent Renewable Portfolio Standard," Office of Integrated Analysis and Forecasting of the Energy Information Administration; Chen, Cliff, Ryan Wiser, and Mark Bolinger, 2007, "Weighing the Costs and Benefits of State Renewables Portfolio Standards: A Comparative Analysis of State-Level Policy Impact Projections," Report prepared for the Permitting, Siting, and Analysis Division, Office of Electricity Delivery and Energy Reliability, U.S. Department of Energy.

[2]A total of 885 respondents completed the first wave. Data were collected by a survey research company (Bovitz Research Group of Encino, CA) and, as with most internet survey samples, respondents participate in multiple surveys over time and receive compensation for their participation. Respondents had a median age of 49, and were 49.0% female, 75.0% white, and 98.9% had at least high school degrees and 80.9% had university degrees. The partisan composition was 39.2% Democrats and 30.7% Republicans.

[3]Given that the issue received almost no media coverage in the intervening months, there is no reason to believe that any subjects were substantially influenced between the earlier and present wave (an assumption that is testable by comparing present attitudes of those in the control condition against their responses in the earlier wave). On average, opinions became slightly more negative between the two waves.

Figure 1: Experimental Design and Treatment Group Sample Sizes



Differences in sample sizes within the captive conditions reflect random assignment. Differences in sample sizes between "Chose Pro" and "Chose Con" conditions reflect treatment self-selection. Choice conditions were intentionally oversampled.

an instrument for respondents' information choices (which would otherwise be nonrandomly determined) and provides insight into the extent to which self-selection behavior varies across contexts. The logic of the manipulation was that individuals who believe their self-interest is at-stake would display higher attitude importance and thus be more likely to choose a treatment message (if given the choice) that was congruent with their prior attitudes (Krosnick 1989; Boninger, Krosnick, and Berent 1995). Importance was manipulated to be high by telling respondents: "A new law is currently moving through Congress that would require your electricity provider to purchase energy from renewable sources (e.g. wind and solar). This is relevant to you since it will influence your energy bills and the environment. The law would go into effect immediately." Those in low importance condition read that "Some have proposed a bill that would require electricity providers to purchase energy from renewable sources (e.g. wind and solar). This is probably not directly relevant to you because Congress does not appear to be ready to act on the bill and even if they did it is unlikely to personally affect you."[4]

The second manipulation varied whether the information respondents read was supportive (Pro) or opposed (Con) to the policy. The Pro message was entitled "Renewable Energy Rules Beneficial" and the Con message was entitled "Renewable Energy Rules Ineffective." Exact text of the messages is available in Appendix A.[5]

[4]Three rounds of pretests were conducted with 80 participants each in order to determine the question and manipulation wordings. Participants were recruited from Amazon Mechanical Turk and compensated $.20 each for their participation. In the final pretest, the high importance manipulation achieved a mean importance (on a 7-point scale) of 5.66 (SD=1.08), while the low importance manipulation achieved a mean importance of 5.2 (SD=1.44). This effect is small, but in the intended direction. It is also a larger effect than either of the previous pretests, which is indicative of the difficulty of manipulating attitude importance. A post-treatment manipulation check asked "How important to you personally is your opinion about this renewable energy restriction?" The results confirms that attitude importance was manipulated. On a 0-1 scale, those in high importance conditions rated their attitude importance at 0.77 and those in low importance conditions rated their attitude importance at 0.69, a statistically significant difference ($p = 0.00$). The mean attitude importance reported by a control condition that received no manipulation and no information was 0.72. There was also no indication that the manipulation of importance had a persuasive effect on opinions of those in the high versus low conditions (p=.42).

[5]An effort was made to ensure that the informational content of the Pro and Con messages was near-identical. The difference between the two treatments is in the language chosen to describe the same basic facts. Pretesting, describe in previous note, confirmed that messages were equally effective. The Pro message was rated as 5.08 (SD=1.38) on a 7-point effectiveness scale, while the Con message was rated 4.97 (SD=1.47) on the same scale. And, the Pro message produced attitudes (on a 7-point scale) of 5.63 (SD=1.64) and

The final manipulation involved how the informational treatments were assigned to respondents. One-third of the respondents were assigned to "captive exposure" conditions that involved simply reading either the Pro or Con argument (based on random assignment). The other two-thirds were presented with the headlines for each passage and told to choose one of them to read, which they were then given, directly emulating the process of selective exposure. When examining the choices made by these respondents, it is clear that the manipulation of attitude importance was effective: 88% of high importance respondents chose information congruent with their $t1$ opinion, while only 63% of low importance respondents chose information congruent with their $t1$ opinion (a statistically significant difference). This means the odds of a high importance respondent choosing the message congruent with their $t1$ opinion was roughly 7:1, while the odds for a low importance subject were only about 1.5:1, yielding a odds-ratio of 4.25, which is a dramatic and significant effect.[6]

After receiving the importance and informational manipulations, outcome measures were collected including evaluations of the information received, attitude toward the policy, subjective intentions to obtain further information, and behavioral measure directly tapping willingness to receive additional information in the form of an email message. All variables are coded to range from 0 to 1, with higher scores indicating more positive evaluations, higher policy support, or greater intention to seek information. The opinion question read,

the Con message produced attitudes at 4.5 (SD=1.93), a difference that is in the intended direction. The difficulty of manipulating attitude importance Visser, Holbrook, and Krosnick (2007) makes the relatively small apparent effects in the small-n pretest seem reasonable.

[6]The odds-ratio would be equal to one if there was no effect of importance. In this case, the 95% confidence interval for the odds-ratio is (2.68,6.75), suggesting a large effect of attitude importance on attitude-congruent selective exposure. This result also holds when looking at supporters and opponents separately: among those with high importance attitudes, 88% of supporters chose the Pro information and 85% of opponents chose the Con information. By contrast, under low importance, only 65% of supporters chose the Pro information and a mere 35% of opponents chose the Con information. Given that all respondents were presented with the two headlines in the same order — the Pro headline coming first — it is possible that the slight preference for the Pro information among those with low importance is attributable to primacy effects. It is also possible that low importance respondents might have opted to read nothing if they had been given the choice, meaning that lacking importance people disengage from political information entirely. Furthermore, given that respondents' opinions leaned positive ($\bar{x}_{t1}$=0.76, SD=0.01), breaking out the results in this way confirms that the aggregate rates of attitude-congruent selective exposure are not simply a result of the sample disproportionately choosing the Pro message. As a comparison, of those with neutral $t1$ opinions, 49% chose the Con message and 51% chose the Pro message, suggesting their choices were relatively arbitrary and confirming the pretest finding that neither the Pro headline or Con headline was more enticing to read.

"Thinking about energy related restrictions, to what extent do you oppose or support requiring electricity providers to purchase energy generated from renewable sources (e.g., wind, solar)?" and solicited responses on a seven-point scale from "strongly opposed" to "strongly support." The argument evaluation question read, "How effective would you say the information you read was in making an argument about this energy-related restriction?" The subjective information-seeking question asked "How likely are you to seek more information about renewable energy requirements?" The behavioral measure asked "Can we send you an email with more information about renewable energy requirements?" Responses to the latter measure were coded as 1 if the respondent entered their email address and 0 otherwise. The two information-seeking measures correlate to some extent ($r = 0.44$). Treatment group means for all outcome measures are reported in Appendix B.

Following from Gaines and Kuklinski (2011*a*), I estimate three different treatment effects for each outcome variable. The first is the familiar sample average treatment effect (SATE):

$$SATE = \bar{Y}_T - \bar{Y}_C \tag{1}$$

where $\bar{Y}_T$ is the mean outcome value among those captively assigned to the Pro message and $\bar{Y}_C$ is the mean outcome value among those captively assigned to the Con message. As Gaines and Kuklinski point out, this SATE is a weighted average of two effects for different observable subsamples: one for those who choose the treatment and one for those who do not (i.e., those who would choose the Pro message and those who would choose the Con message in this particular experiment; I ignore the true control group in these calculations). The second estimate is therefore the effect of the treatment on those who would choose it (the *Pro-selector effect* of the Pro message):

$$T_s = \frac{\bar{Y}_S - \bar{Y}_C}{\hat{\alpha}} \tag{2}$$

where $\bar{Y}_S$ is the mean outcome value among all respondents assigned to the "choice" condition

(combining those who chose the Pro and Con messages) and $\hat{\alpha}$ is the proportion of these individuals choosing the Pro message. The third estimate is the analogous effect among those who would not choose the treatment (i.e., the *Con-selector effect* of the Pro message):

$$T_n = \frac{\bar{Y}_T - \bar{Y}_S}{1 - \hat{\alpha}} \tag{3}$$

The latter two effects are causally identified given random assignment to the choice and captive arms of the experiment. It is important to note that in the present study these are not the effects of receiving the Pro and Con messages, but rather the effect of receiving the Pro message on different subsets of the sample that differ in their preferences over the available treatment alternatives. In other words, the Pro-selector effect represents the effect of the Pro message treatment for those who would opt for the Pro message *if given the opportunity* and the Con-selector effect represents the effect of the Pro message treatment for those who would opt for the Con message *if given the opportunity*.

I further report each effect estimate for the sample as a whole, as well as separately for those in the low importance and high importance conditions. Separately estimating the effects in this way makes it possible to detect further effect heterogeneity when the pattern of treatment self-selection changes.

Results

Table 1 presents the main results, separately for the full sample in panel (a), the low importance condition in panel (b), and the high importance condition in panel (c). Looking at panel (a), we see sample average treatment effects (SATEs) estimated from the captive conditions for each of the five outcome measures (argument evaluation, opinion level, opinion change, planned information-seeking, and requests for an email with issue-relevant information), alongside estimated effects for Pro-selectors (T_s) and Con-selectors (T_n). In substantive terms, these five SATEs indicate: (1) the Pro argument is seen as more effective than the

Con argument, which makes sense given the supportive attitudes of the sample, (2) exposure to the Pro message increases support for the policy, (3) this increase in support holds when measured by $t2 - t1$ changes in opinion rather than $t2$ levels of opinion, (4) the Pro message insignificantly reduces self-reported plans to seek out issue-relevant information, and (5) the Pro message reduces the likelihood of requesting an email with more policy information.

The Pro-selector effects (T_s) paint a largely similar story, with differences only in the magnitude but not direction of effects. The story is different for Con-selectors: while the Pro message still increases policy support (as measured by changes over time), there are no other substantively sizable effects on exposure and none of the T_n effects are statistically distinguishable from zero. In other words, there appears to be effect heterogeneity: the Pro-selectors are affected by the Pro message in various ways, while there appears to be no effect of the message for Con-selectors (i.e, those who would avoid it if given the chance).

Turning to panel (b) of Table 1, which displays the low importance condition, the SATEs are very similar to those for the sample as a whole: the Pro message is seen as more effective, it increases policy support as measured by both $t2$ opinions and opinion changes over time, and decreases the rate of requesting a follow-up email, while there is no effect on intended information-seeking. Turning to the Pro-selector effects in column 2, we see a pattern generally consistent with the SATEs in effect direction but only one effect is statistically distinguishable from zero: exposure to the Pro message leads to a substantial decrease in requests for the follow-up email. The Con-selector effects in the third column are also consistent with the SATEs in direction but only the effect on opinion changes (wherein exposure to the Pro message increases support) is statistically distinguishable from zero. Under low importance, the there appears to be relatively little effect heterogeneity: effects for Pro-selectors and Con-selectors are fairly consistent with one another, meaning the SATEs provide inferences that seem to apply equally well to those preferring and not preferring the treatment.

Finally, panel (c) of Table 1 shows results for the high importance condition. Recall

Table 1: Treatment Effect Estimates, by Importance Condition

	SATE	T_s	T_n
Evaluation	0.14 (0.03)	0.21 (0.04)	-0.01 (0.08)
Opinion	0.08 (0.04)	0.12 (0.04)	0.01 (0.09)
Opinion Change ($t2 - t1$)	0.11 (0.03)	0.12 (0.04)	0.09 (0.08)
Information-Seeking	-0.02 (0.03)	-0.03 (0.04)	0.02 (0.08)
Email Request	-0.13 (0.06)	-0.20 (0.07)	0.04 (0.15)

(a) Full sample

	SATE	T_s	T_n
Evaluation	0.14 (0.02)	0.09 (0.05)	0.11 (0.09)
Opinion	0.14 (0.03)	0.04 (0.06)	0.07 (0.11)
Opinion Change ($t2 - t1$)	0.08 (0.03)	0.06 (0.06)	0.20 (0.10)
Information-Seeking	0.03 (0.03)	-0.05 (0.06)	0.05 (0.10)
Email Request	-0.12 (0.05)	-0.32 (0.10)	-0.17 (0.17)

(b) Low Importance Condition

	SATE	T_s	T_n
Evaluation	0.21 (0.03)	0.31 (0.05)	-0.20 (0.14)
Opinion	0.31 (0.03)	0.18 (0.06)	-0.08 (0.17)
Opinion Change ($t2 - t1$)	0.08 (0.03)	0.17 (0.05)	-0.07 (0.13)
Information-Seeking	0.05 (0.03)	-0.01 (0.04)	-0.03 (0.13)
Email Request	0.05 (0.05)	-0.09 (0.09)	0.36 (0.28)

(c) High Importance Condition

Note: Cell entries are estimated treatment effects, with bootstrapped standard errors in parentheses. The estimated values of $\hat{\alpha}$ used in estimating T_s and T_n are: 0.69 (full sample), 0.63 (low importance), and 0.75 (high importance)

that under high importance, respondents were much more likely to self-select a treatment message congruent with their prior ($t1$) opinion than in the low importance condition (the odds-ratio was 4.25). The consequence of this for inferences about the effects of the treatment messages should be immediately clear: while the SATEs in this condition mirror those for the low importance condition and the sample as a whole (with the exception of the email message outcome), the effects for Pro-selectors and Con-selectors come out substantially different. Under high importance, there is a very large positive Pro-selector effect on argument evaluation (meaning the Pro message was seen much more favorably than the Con message) and opinions were moved nearly 20% more supportive among Pro-selectors. The Pro-selector effects for the two information-seeking measures were not distinguishable from zero. Looking at the Con-selector effects in column 3, the effects on argument evaluation and opinion now are substantively negative, which indicates that the Pro message was seen as less effective than the Con message and actually produced a backlash that made these respondents less supportive of the policy, though the large standard errors mean these effects are not statistically distinguishable from zero. Looking at the information-seeking measures, there is a substantively large positive effect of the Pro message on email requests but again the large standard errors mean this is difficult to distinguish from statistical noise. The high importance conditions therefore suggest a quite different pattern of causal effects than the low importance conditions, in both size and direction of effects for the Pro-selectors and Con-selectors.

These results are striking. The subsample effects vary widely between the high and low importance conditions. If one were to use these data to make inferences about the effects of exposure to a political argument, those inferences would differ substantially depending whether one looked only at the SATEs or at effects derived from the hybrid design, and those effects would differ even further depending on the level of attitude-congruent selective exposure driving the participants' choices of message. Similarly interesting is the fact that the SATEs are very similar across these conditions despite the underlying group-specific

effects differing considerably. This means that the randomized experiment offers a very similar inference about the effects of exposure to a political communication even when the underlying group-specific effects for individuals who prefer the Pro and Con messages are quite different. While social scientists frequently acknowledge that the SATE is necessarily an average of underlying individual-level effects and may not represent the effect for any given unit or group of units, these results show that the SATE can be quite misleading. Yet the effects for Pro-selectors and Con-selectors (or, more generically, treatment choosers and avoiders) may be similarly misleading to the extent the size and direction of these effects depended heavily on the pattern of treatment self-selection determined by the importance manipulation. This means that a self-selection component of an experiment presents a face valid form of experimental realism, but that realism does not necessarily lead to clearer inferences about the effects of a treatment that is prone to self-selection.

Discussion

The present experiment examined how treatment self-selection can be used to study political phenomena characterized by real-world choices, namely self-selected exposure to political communications (Bennett and Iyengar 2008; Lau and Redlawsk 2006; Levendusky 2012; Arceneaux and Johnson 2012; Stroud 2011). Building on Gaines and Kuklinski's (2011*a*) hybrid experimental design, the results show that an experiment randomly assigning participants to receive different messages appears to mask substantial effect heterogeneity among those that would be inclined to select distinct messages (i.e., self-select into the various treatment alternatives). Moving beyond previous work, these results additionally demonstrate that such heterogeneity is quite complex: individuals' choices are themselves shaped by various factors, including the perceived importance of a policy issue.

As such, providing an opportunity for treatment self-selection inside an experimental context can reveal useful information about the heterogeneity in individuals' responses to

particular treatments. But that self-selection component is no panacea for ecological validity concerns. When researchers believe that treatment effects are likely to vary across individuals with different preferences for exposure to the various treatment alternatives, it is likely insufficient to simply conduct a hybrid experiment without further consideration about who will select which treatment(s) and why. Treatment self-selection is itself a complex process that — in many domains — is unlikely to be well-summarized by a single choice condition and no single strategy for incorporating choice into an experiment is likely to provide perfect ecological validity (see, for example, Feldman et al. 2013).

Experimental studies of political engagement, selective exposure, and motivated thinking should explore the democratic consequences of message self-selection (and the attitude importance that drives it) more thoroughly. The results here suggest that randomized exposure and self-selected exposure to a brief political message seem to produce distinct inferences about the effects of that exposure. The self-selection conditions offer a uniquely realistic degree of ecological validity, but with that realism comes new challenges. While the findings here suggest incorporating treatment self-selection into a randomized experiment can be quite fruitful, more research is needed to understand how best to study self-selection processes using experimental methods.

References

Ansolabehere, Stephen, Shanto Iyengar, Adam F. Simon, and Nicholas A. Valentino. 1994. "Does Attack Advertising Demobilize the Electorate?" *American Political Science Review* 88(4): 829–838.

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

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.

Berinsky, Adam J., and Donald R. Kinder. 2006. "Making Sense of Issues Through Media Frames: Understanding the Kosovo Crisis." *The Journal of Politics* 68(3): 640–656.

Bolsen, Toby, and Thomas J. Leeper. 2013. "Self-Interest and Attention to News among

Issue Publics." *Political Communication* In press(3): 329–348.

Boninger, David S., Jon A. Krosnick, and Matthew K. Berent. 1995. "Origins of Attitude Importance: Self-Interest, Social Identification, and Value Relevance." *Journal of Personality and Social Psychology* 68(1): 61–80.

Brewer, Paul R. 2003. "Values, Political Knowledge, and Public Opinion About Gay Rights: A Framing-Based Account." *Public Opinion Quarterly* 67(2): 173–201.

Brewer, Paul R., and Kimberly Gross. 2005. "Values, Framing, and Citizens' Thoughts about Policy Issues: Effects on Content and Quantity." *Political Psychology* 26(6): 929–948.

Chong, Dennis, and James N. Druckman. 2007. "Framing Public Opinion in Competitive Democracies." *American Political Science Review* 101(4): 637–655.

Disch, Lisa. 2011. "Toward a Mobilization Conception of Democratic Representation." *American Political Science Review* 105(01): 100–114.

Ditto, Peter H., James A. Scepansky, Geoffrey D. Munro, Anne Marie Apanovitch, and Lisa K. Lockhart. 1998. "Motivated Sensitivity to Preference Inconsistent Information." *Journal of Personality and Social Psychology* 75(1): 53–69.

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.

Feldman, Lauren, Edward W. Maibach, Connie Roser-Renouf, and Anthony A. Leiserowitz. 2011. "Climate on Cable: The Nature and Impact of Global Warming Coverage on Fox News, CNN, and MSNBC." *The International Journal of Press/Politics* , 2–31.

Feldman, Lauren, Natalie Jomini Stroud, Bruce Bimber, and Magdalena E. Wojcieszak. 2013. "Assessing Selective Exposure in Experiments: The Implications of Different Methodological Choices." *Communication Methods and Measures* (August): 37–41.

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

Gaines, Brian J., and James H. Kuklinski. 2011*b*. "Treatment Effects." In *Cambridge Handbook of Experimental Political Science*, eds. James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia. New York: Cambridge University Press , 445–458.

Garrett, R. Kelly. 2009*a*. "Echo Chambers Online? Politically Motivated Selective Exposure Among Internet News Users." *Journal of Computer-Mediated Communication* 14(2): 265–285.

Garrett, R. Kelly. 2009*b*. "Politically Motivated Reinforcement Seeking: Reframing the Selective Exposure Debate." *Journal of Communication* 59(4): 676–699.

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.

Hart, William, Dolores Albarracín, Alice H. Eagly, Inge Brechan, Matthew J Lindberg, and Lisa Merrill. 2009. "Feeling Validated Versus Being Correct: A Meta-Analysis of Selective Exposure to Information." *Psychological Bulletin* 135(4): 555–588.

Holbrook, Allyson L., Matthew K. Berent, Jon A. Krosnick, Penny S. Visser, and David S. Boninger. 2005. "Attitude Importance and the Accumulation of Attitude-Relevant Knowledge in Memory." *Journal of Personality and Social Psychology* 88(5): 749–769.

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

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

Iyengar, Shanto, and Kyu S. Hahn. 2009. "Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use." *Journal of Communication* 59(1): 19–39.

Iyengar, Shanto, Kyu S. Hahn, Jon A. Krosnick, and John Walker. 2008. "Selective Exposure to Campaign Communication: The Role of Anticipated Agreement and Issue Public Membership." *The Journal of Politics* 70(01): 186–200.

Kim, Young Mie. 2007. "How Intrinsic and Extrinsic Motivations Interact in Selectivity: Investigating the Moderating Effects of Situational Information Processing Goals in Issue Publics' Web Behavior." *Communication Research* 34(2): 185–211.

Kim, Young Mie. 2009. "Issue Publics in the New Information Environment: Selectivity, Domain Specificity, and Extremity." *Communication Research* 36(2): 254–284.

Krosnick, Jon A. 1989. "Attitude Importance and Attitude Accessibility." *Personality and Social Psychology Bulletin* 15(3): 297–308.

Krosnick, Jon A. 1990. "Government Policy and Citizen Passion: A Study of Issue Publics in Contemporary America." *Political Behavior* 12(1): 59–92.

Kunda, Ziva. 1990. "The Case for Motivated Reasoning." *Psychological Bulletin* 108(3): 480–498.

Lau, Richard R., and David P. Redlawsk. 2006. *How Voters Decide: Information Processing in Election Campaigns*. New York: Cambridge University Press.

Levendusky, Matthew S. 2012. *Partisan News that Matters: How Partisan News Impacts American Politics*. Chicago, IL: The University of Chicago Press.

Levendusky, Matthew S. 2013. "Why Do Partisan Media Polarize Viewers?" *American Journal of Political Science* xx(xx): no–no.

Lodge, Milton, and Charles S. Taber. 2000. "Three Steps Toward a Theory of Motivated Political Reasoning." In *Elements of Reason: Cognition, Choice, and the Bounds of Rationality*, eds. Arthur Lupia, Mathew D. McCubbins, and Samuel L. Popkin. New York: Cambridge University Press , 183–213.

McDermott, Rose. 2011. "Internal and External Validity." In *Cambridge Handbook of Exper-*

imental Political Science, eds. James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia. New York: Cambridge University Press.

Miller, Joanne M., and Jon A. Krosnick. 2000. "News Media Impact on the Ingredients of Presidential Evaluations: Politically Knowledgeable Citizens are Guided by a Trusted Source." *American Journal of Political Science* 44(2): 301–315.

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.

Nickerson, Raymond S. 1998. "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises." *Review of General Psychology* 2(2): 175–220.

Petty, Richard E., and John T. Cacioppo. 1986. "The Elaboration Likelihood Model of Persuasion." *Advances in Experimental Social Psychology* 19: 123–205.

Sears, David O., and Jonathan L. Freedman. 1967. "Selective Exposure to Information: A Critical Review." *Public Opinion Quarterly* 31(2): 194–213.

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

Smith, Steven M., Leandre R. Fabrigar, and Meghan E. Norris. 2008. "Reflecting on Six Decades of Selective Exposure Research: Progress, Challenges, and Opportunities." *Social and Personality Psychology Compass* 2(1): 464–493.

Steiner, Peter M., Thomas D. Cook, William R. Shadish, and M.H. Clark. 2010. "The Importance of Covariate Selection in Controlling for Selection Bias in Observational Studies." *Psychological Methods* 15(3): 250–267.

Stroud, Natalie Jomini. 2011. *Niche News: The Politics of News Choice*. New York: Oxford University Press.

Taber, Charles S., and Milton Lodge. 2006. "Motivated Skepticism in the Evaluation of Political Beliefs." *American Journal of Political Science* 50(3): 755–769.

Tichenor, P.J., G.A. Donohue, and C.N. Olien. 1970. "Mass Media Flow and Differential Growth in Knowledge." *Public Opinion Quarterly* 34(2): 159–170.

Visser, Penny S., Allyson L. Holbrook, and Jon A. Krosnick. 2007. "Knowledge and Attitudes." In *Public Opinion Research (SAGE Handbook)*. Thousand Oaks, CA: Sage Publications , 123–140.

Appendix A. Survey Material

Attitude Importance Manipulation:
High: A new law is currently moving through Congress that would require your electricity provider to purchase energy from renewable sources (e.g. wind and solar). This is relevant to you since it will influence your energy bills and the environment. The law would go into effect immediately.
Low: A few legislators in Congress have proposed a bill that would require electricity providers to purchase energy from renewable sources (e.g. wind and solar). This is probably not directly relevant to you because Congress does not appear to be ready to act on the bill and even if they did it is unlikely to personally affect you.

Information Choice Manipulation:
Assigned: [Randomly assigned to Pro or Con]

Choice: Before you proceed, please choose one of the following brief passages to read:
"Renewable Energy Rules Beneficial"
"Renewable Energy Rules Ineffective"

Article Text:
Pro: "Renewable Energy Rules Beneficial"
The proposed federal law would create uniform nationwide standards which is necessary since many states have not adopted renewable energy provisions. The new standards would require electricity utilities to produce between 10% and 30% of their energy from renewable sources especially wind power as well as potentially innovative new sources of energy. This in turn would reduce pollution. The impact on consumers is also affordable: adopting a nationwide standard would increase monthly electricity bills by only about 1%. Renewable standards therefore reduce reliance on fossil fuels for energy production without dramatically increasing costs to American consumers.

Con: "Renewable Energy Rules Ineffective"
The proposed federal law would intervene in state policies and regulate private businesses to create uniform nationwide standards, where up until now many states have not adopted renewable energy provisions. The new standards would drive down innovation by requiring energy utilities to adopt specific technologies (e.g. wind power) rather than directly targeting the reduction of polluting greenhouse gas emissions. The impact on consumers is also problematic: adopting a nationwide standard would increase monthly electricity bills by up as much as 4%. Renewable standards therefore increase the cost of energy through government regulation without directly addressing potential environmental impacts of energy production from fossil fuels.

Appendix B. Supplementary Descriptive Statistics

Table B1a. Argument Evaluations, by Treatment Condition

	Mean (SE)	N
High Choice Pro	0.75 (0.01)	204
High Choice Con	0.57 (0.03)	67
High Captive Pro	0.65 (0.03)	66
High Captive Con	0.47 (0.04)	68
Low Choice Pro	0.69 (0.02)	173
Low Choice Con	0.53 (0.03)	101
Low Captive Pro	0.68 (0.03)	67
Low Captive Con	0.58 (0.03)	67

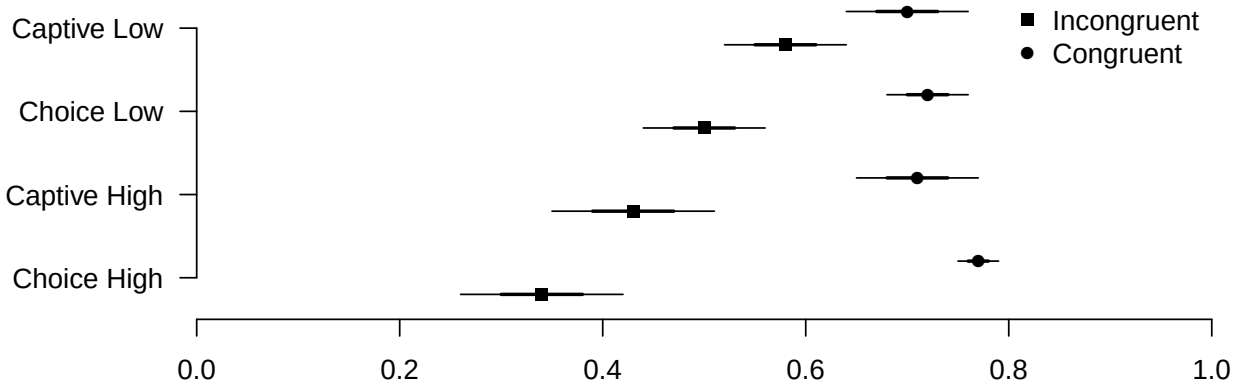
Note: Cell entries are treatment group means with standard errors in parentheses and higher values indicating greater perceived argument effectiveness (scaled 0-1).

Table B1b. Argument Evaluations, by Importance and Choice Condition

	Mean (SE)	N
Choice High	0.70 (0.02)	271
Captive High	0.56 (0.03)	134
Choice Low	0.63 (0.02)	274
Captive Low	0.63 (0.02)	134

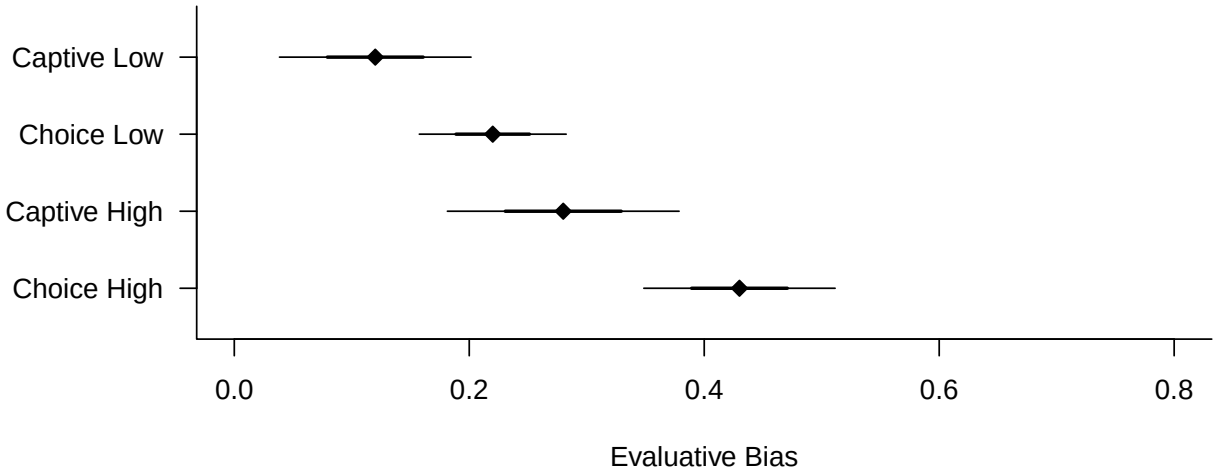
Note: Cell entries are treatment group means with standard errors in parentheses and higher values indicating greater perceived argument effectiveness (scaled 0-1).

Figure B1. Mean Argument Evaluations, by Importance and Choice Condition and Congruence with $t1$ Opinion



Note: Figure displays mean argument evaluations (and bars representing one and two standard errors of the mean) by importance and choice Condition, separately for arguments congruent or incongruent with respondents' $t1$ opinions.

Figure B2. Argument Evaluation Bias, by Importance and Choice Condition



Note: Figure displays the bias toward seeing attitude-congruent messages as more effective than attitude-incongruent messages. Points represent difference-in-differences estimates along with bars representing one and two associated standard errors, based on the mean argument ratings displayed in Figure C1.

Table B2a. $t1$ **and** $t2$ **Opinions, by Treatment Condition**

	t1 Mean (SE)	t2 Mean (SE)	t2-t1 Mean (SE)	N
Control	0.78 (0.03)	0.65 (0.03)	-0.14 (0.03)	66
High Choice Pro	0.86 (0.01)	0.82 (0.01)	-0.04 (0.02)	204
High Choice Con	0.47 (0.04)	0.39 (0.04)	-0.08 (0.04)	67
High Captive Pro	0.76 (0.03)	0.70 (0.04)	-0.07 (0.03)	66
High Captive Con	0.75 (0.04)	0.58 (0.04)	-0.18 (0.04)	68
Low Choice Pro	0.82 (0.02)	0.74 (0.02)	-0.07 (0.02)	173
Low Choice Con	0.69 (0.02)	0.55 (0.03)	-0.14 (0.03)	101
Low Captive Pro	0.72 (0.04)	0.70 (0.04)	-0.02 (0.03)	67
Low Captive Con	0.78 (0.03)	0.65 (0.03)	-0.13 (0.03)	67

Note: Cell entries are treatment group means with standard errors in parentheses and higher values indicating greater support (scaled 0-1).

Table B2b. $t1$ **and** $t2$ **Opinions, by Importance and Choice Condition**

	t1 Mean (SE)	t2 Mean (SE)	t2-t1 Mean (SE)	N
Choice High	0.76 (0.02)	0.72 (0.02)	-0.05 (0.02)	271
Captive High	0.76 (0.02)	0.64 (0.03)	-0.12 (0.02)	134
Choice Low	0.77 (0.02)	0.67 (0.02)	-0.10 (0.02)	274
Captive Low	0.75 (0.02)	0.67 (0.03)	-0.08 (0.02)	134

Note: Cell entries are treatment group means with standard errors in parentheses and higher values indicating greater support (scaled 0-1).

Table B3a. Information Seeking, by Treatment Condition

	Subjective	Email	N
Control	0.62 (0.03)	0.41 (0.07)	66
High Choice Pro	0.68 (0.02)	0.54 (0.03)	204
High Choice Con	0.59 (0.04)	0.43 (0.07)	67
High Captive Pro	0.65 (0.03)	0.61 (0.05)	66
High Captive Con	0.67 (0.03)	0.59 (0.05)	68
Low Choice Pro	0.54 (0.02)	0.31 (0.05)	173
Low Choice Con	0.49 (0.03)	0.34 (0.07)	101
Low Captive Pro	0.54 (0.03)	0.25 (0.09)	67
Low Captive Con	0.56 (0.03)	0.52 (0.06)	67

Note: Cell entries are treatment group means with standard errors in parentheses and higher values indicating greater information (scaled 0-1) and likelihood of requesting email (0/1).

Table B3b. Information Seeking, by Importance and Choice Condition

	Subjective	Email	N
Choice High	0.66 (0.01)	0.52 (0.03)	271
Captive High	0.66 (0.02)	0.60 (0.03)	134
Choice Low	0.52 (0.02)	0.32 (0.04)	274
Captive Low	0.55 (0.02)	0.39 (0.05)	134

Note: Cell entries are treatment group means with standard errors in parentheses and higher values indicating greater information (scaled 0-1) and likelihood of requesting email (0/1).